ON TRUNCATED SEQUENTIAL TESTS OF PARAMETERS IN VARIOUS POISSON MODELS WITH APPLICATIONS TO TRAFFIC ACCIDENTS

Stig Elofsson

Stockholm 1975

ON TRUNCATED SEQUENTIAL TESTS OF PARAMETERS IN VARIOUS POISSON MODELS WITH APPLICATIONS TO TRAFFIC ACCIDENTS

Stig Elofsson

Stockholm 1975

ISBN 91–7222–130–5
Sundt & Co Stockholm

To
GUNILLA
ARNE ERIKA PATRIK

UNIV.
LEIDEN
BIBL.

PREFACE

The purpose of this study is to investigate sequential test procedures — especially truncated tests — and to compare them with fixed-sample size tests.

Suppose the observations $x_1, x_2 \ldots$ follow a Poisson distribution with probability function $f(x;\lambda)$. Then, according to Wald (1947), a sequential probability ratio test (SPRT) of $H_0:\lambda \leq \lambda_0$ against $H_1:\lambda \geq \lambda_1$ can be constructed in the following way. As long as the cumulative log likelihood ratio

$$S_n = \sum_1^n z_i = \sum_1^n \ln \{\frac{f(x;\lambda_1)}{f(x;\lambda_0)}\}$$

lies between two constants a and b, determined by the chosen risk levels, continue sampling. Terminate the test as soon as $S_n \leq b$ — in which case H_0 is accepted — or $S_n \geq a$ — in which case H_1 is accepted. The number of observations necessary to terminate the test is a random variable and is called the decisive sampling number, N. Wald showed that a test could be constructed which minimizes the average sampling number when λ equals λ_0 or λ_1. However, this test has some shortcomings. The average sampling number can be high for some values of λ between λ_0 and λ_1, the variance of N can be large and the distribution of N has positive skewness. Some of these drawbacks can be overcome by truncating the test at N_0, i.e. the sampling is stopped after N_0 observations have been made.

Under the assumption that the log likelihood ratio, Z, is normally distributed, Wald derived an approximate analytic expression of the characteristic function of N. From this expression, Anderson (1960) and Gosh (1970) derived expressions for the operating characteristic (OC) and average sampling number (ASN) functions of truncated sequential probability ratio tests.

Under the assumption that the observations follow a Poisson distribution, Z will not be normally distributed. In this study the actual distribution of Z has been approximated with a normal distribution. First, the parameter of this normal distribution were chosen equal to the mean and the variance of Z, but this approximation turned out to be too crude. According to Wald's derivation, an analytic expression of the non-zero root h_0 to the equation $\ln\varphi_\lambda(h) = 0$ is used, where $\varphi_\lambda(h)$ is the moment-generating function of Z. The parameters of the approximate distribution of Z are directly related to the expression of h_0, so that the problem entails finding a suitable analytic approximation of h_0. Boundaries to the sampling continuation region for S_n, as well as the truncation point in truncated tests, are determined by the distribution of N at λ equal to λ_0 and λ_1. The analytic expression of h_0 then has to be chosen so that it gives the correct values of h_0 at $\lambda = \lambda_0$ and λ_1.

The approximations made imply that the discrete Poisson process is approximated by a continuous Wiener-process. The approximations are evaluated in terms of the central moments of the decisive sampling number, N, in the SPRT and of risk levels and average sampling number in truncated tests.

It is shown that the approximation of the discrete process to a continuous process introduces larger errors in the approximations of the average sampling number than the approximation of the Poisson distribution with a normal distribution. The average sampling number in the tests will be underestimated and the underestimate increases with λ_0 and the ratio λ_1/λ_0.

Different test procedures are compared and it is found that the sequential test procedures are better than fixed-sample size tests. One-sided truncated SPRT's can be used in special situations, when either large values of λ or small values of λ are relevant. A two-sided truncated SPRT will reduce the variance and the coefficient of skewness of the decisive sampling number, especially for values of λ close to $(\lambda_1+\lambda_0)/2$, as compared to the SPRT.

Various stochastic models, where the original observations are assumed to be Poisson distributed, have been treated. The effect of the heterogeneity in the values of λ, i.e. the means of the observations vary, has been studied through the negative binomial distribution. This heteregoneity has little effect on the OC-function, whereas the effect on the ASN-function is greater (higher variance in the means increasing the average sampling number).

The practical use of sequential tests gives rise to problems about the choice of an appropriate stochastic model, risk levels, hypotheses, etc. Examples of applications of sequential tests in the analysis of road traffic accidents are presented in Chapter VII in order to point out these problems and possible solutions to them.

** ******

I wish to express my thanks to the following people for their support and assistance. Professor Sten Malmquist supervised my research, read through all the "versions" and helped me to transform them into this book. Docent Harry Lütjohann gave me valuable comments in the initial stage of the research. Ulf Thelander read the first version of the manuscript and his comments have improved this book. Julie Sundqvist not only checked my English but also helped me to improve the form of the manuscript. Earlier versions of the manuscript were typed by Kari Blomberg, Anna-Lisa Persevall and Elisabeth Sundin. The final text was typed by Marion Ekström and Eva Rudh. Finally, I would like to thank the Transport Research Delegation for financially supporting the research, which forms the basis of this monograph.

Stockholm, May 1975

Stig Elofsson

CONTENTS

I	Background and results	
I.1	Introduction	7
I.2	Statistical models and sampling models for traffic accidents	8
I.3	Sequential tests and their efficiency	9
I.4	Derivation and comparison of tests for the one-sample case	12
I.5	Comparing the parameters of two homogeneous Poisson distributions	17
I.6	Heterogeneity in parameter values	18
II	Approximations of the distribution of the sample size for a sequential probability ratio test of the Poisson mean	
II.1	Introduction	23
II.2	Analytic expressions of the roots to the equation $\ln\varphi(h) = -it$	25
II.3	An approximation to the distribution of N	31
III	Truncated sequential tests	
III.1	Generalized SPRT and other optimal tests procedures	36
III.2	Truncated SPRT of Poisson parameters	40
III.3	Sequential probability ratio tests for comparing the means of two Poisson distributed random variables	46
III.3.1	Matched observations	46
III.3.2	Comparison of a fixed-sample with sequential observations	51
III.4	Sequential probability ratio tests of the parameters in a negative binomial distribution	62
III.5	Sequential probability ratio tests of the parameter in a multiplicative Poisson model	64
IV	Evaluation of approximate expressions of test characteristics	
IV.1	Introduction	67
IV.2	Approximation of the discrete state space to a continuous state space	68
IV.3	Error made by approximating the discrete time process to a continuous time process	72
IV.4	Evaluating approximations of characteristics of truncated SPRT	76
V	Comparison of different test procedures	92
VI	Effect of heterogeneity in parameter values	101
VII	On the application of sequential tests in the analysis of road traffic accidents	
VII.1	Introduction	109
VII.2	Fatal accidents in congested areas	110
VII.3	Accidents with personal injury between vehicles and bicycles or mopeds in congested areas	116
	References	124
	Appendix A - F	

Chapter I BACKGROUND AND RESULTS

I.1 Introduction

This study of sequential tests originated from problems encountered in the analysis of road traffic accident data, particularly in connection with the change-over to right hand traffic in Sweden in 1967.

Traffic safety research often concerns studying the effect of a specific treatment on the number of accidents. Many problems are partly or wholy transformed into decision problems involving a choice between two or more decisions. The research in these cases is aimed at making a decision under given risks as quickly as possible. If the number of observed time units, i.e. the time for observing accidents can be reduced, we can

i) better control the experiment situation

ii) study more treatments where the complexity of the experiment plan is given or arrive at less complex experiment plans where the number of treatments is given

iii) lessen the application time of ineffective treatments.

The background to the study of truncated sequential tests is described in this chapter. We turn first to a discussion of the sampling situations and probability models to be studied (Section I.2). Next, Wald's (1947) sequential probability ratio test (SPRT) is presented. The properties of this test are described on the basis of the study of the number of observations necessary to terminate the test, the decisive sampling number, N (Section I.3). An approximation to the distribution of N, presented by Wald (1947), is shown in Section I.4. Results for truncated tests of the parameter in a Poisson distribution are also presented in this section. If the problem is extended to compare two samples, two types of models may be considered. In the first type of model, observations can be matched and treated pairwise, while the other type of model represents a situation where one sample is a fixed-sample and the other contains observations

drawn sequentially (Section I.5). Finally, models which take into account variations between the means of the observations are discussed in Section I.6.

I.2 Statistical Models and Sampling Models for Traffic Accidents

The term sampling models refers to the statistical description (approximation) of the experiment situation. If we study the effect of treatments such as general or differentiated speed limits, safety belts, traffic control, new traffic signs, information, the design of roadway elements, etc) we find that the researcher compares two samples (or series), one for periods or places where the treatment is not applied and the second where the treatment is applied. When using sequential analysis we distinguish between two types of models. In the first type the sampling units of the two samples are treated pairwise. In the second type one of the samples has fixed size and the other sample is studied sequentially.

Erlander (1970) made a survey of the statistical models used to describe the variation in the frequency of traffic accidents. These models are based on the Poisson distribution. Models with a constant accident frequency are usually too simple, so a more complex description of the variation in accident frequency has to be used. One possibility is to introduce several groups (e.g. weekdays) each of which has a constant accident frequency. A model such as this could also be regarded as an approximation to a regression model with characteristics other than time as independent variables (e.g. traffic, weather, roadway characteristics, etc.). Linear regression models of this type have been used in several projects. Another way to allow for variation in the mean accident frequency λ is to let λ be a random variable. If λ follows a gamma distribution, traffic accidents would follow a negative binomial distribution. The model which allows for the greates variation in the mean accident frequency is the multiplicative Poisson model where it is assumed that sampling units with and without the application of treatment can be matched and that the effect of treatment

is proportionally constant. When these assumptions are fulfilled, the mean accident frequency could be different for all sampling units.

The sampling and probability models that will be dealt with are presented in the following outline.

Sampling models

Probability models	One-sample	Comparison of two samples: Pairwise sequential observations	Comparison of two samples: Fixed-sample compared to sequential observations
Poisson distribution, all observations having the same mean	Section III.2 Chapter IV, V	Section III.3.1	Section III.3.2
Negative binomial distribution, i.e. λ is assumed to follow a gamma distribution	Section III.4 Chapter VI		
Multiplicative Poisson model	Section III.5		

I.3 Sequential Tests and Their Efficiency

Wald (1947) presented "the sequential probability radio test", SPRT, which is a parallel to Neyman-Pearson's probability ratio test for fixed sample size with respect to deciding between two simple hypotheses. Suppose a random variable X has a distribution $f(x;\theta)$ and we want to test the hypotheses $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1$. The SPRT decides in favor of either θ_0 or θ_1 on the basis of observations $x_1, x_2, \ldots$ under the following restrictions

$$P(\text{accept } H_0 \mid \theta_0) \geq 1-\alpha$$
$$P(\text{accept } H_0 \mid \theta_1) \leq \beta . \qquad \text{(I.3.1)}$$

For the fixed-sample test, Neyman-Pearson presented the optimal solution for testing H_0 against H_1. Given that the probability of accepting H_0 when θ_0 is true, is larger than or equal to $1-\alpha$, the test is then optimal in the sense that it gives the smallest possible probability of accepting H_0 when θ_1 is true (i.e. it is the most powerful test).

This test depends on the log likelihood ratio

$$S_n = \ln \left\{ \frac{f(x_1 \; x_n; \theta_1)}{f(x_1 \; x_n; \theta_0)} \right\} \quad . \tag{I.3.2}$$

Wald's SPRT is analogous to this test and operates as follows. Continue sampling as long as

$$b < S_n < a \; . \tag{I.3.3}$$

Stop sampling and decide on θ_0 as soon as $S_n \leq b$ and stop sampling and decide on θ_1 as soon as $S_n \geq a$. The constants a and b are chosen so as to fulfill (I.3.1) and Wald showed the probabilities of error α and β will be obtained approximately if the constants are chosen as

$$a = \ln \left[(1-\beta)/\alpha \right] \; ; \; b = \ln \left[\beta/(1-\alpha) \right] \; . \tag{I.3.4}$$

In sequential tests, the number of observations, N, up to and including termination of the test — the decisive sampling number — is a random variable and optimal properties of tests are given by the distribution of N.

The likelihood ratio for many distribution families is a monotonic function of x. This implies that the optimal conditions of tests of the two simple hypotheses could be transformed into tests of $H_0: \theta \leq \theta_0$ against $H_1: \theta \geq \theta_1$. Wald derived several important results for the case where (I.3.2) can be written as

$$S_n = \sum_1^n \ln \left\{ \frac{f(x_i; \theta_1)}{f(x_i; \theta_0)} \right\} = \sum_1^n Z_i \tag{I.3.5}$$

and the Z_i's are identically and independently distributed.

The term operating characteristic function, OC-function, refers to the probability of accepting H_0 when θ is the true parameter value. For this function, Wald showed that

$$Q(\theta) \approx \frac{e^{ah(\theta)} - 1}{e^{ah(\theta)} - e^{bh(\theta)}} \tag{I.3.6}$$

where $h(\theta)$ is the solution to

$$\sum_x \left[\frac{f(x;\theta_1)}{f(x;\theta_0)}\right]^{h(\theta)} f(x;\theta) = 1 \quad \text{or} \quad \int_{-\infty}^{\infty} \left[\frac{f(x;\theta_1)}{f(x;\theta_0)}\right]^{h(\theta)} f(x;\theta)dx = 1 .$$

The optimality of fixed-sample tests is given according to the power of the test. The optimatily of sequential tests is given according to the distribution of N. The average sampling number E(N) as a function of the parameter value θ is called the ASN-function and is ordinarily used in a optimality criterion. Wald derived an approximation to this function, given by

$$E(N;\theta) \approx \frac{Q(\theta) \cdot b + \{1 - Q(\theta)\} \cdot a}{E(z;\theta)} . \qquad (I.3.7)$$

Studies by Baker (1950), Page (1954), Kemp (1958) and Barracbough é Page (1959) indicate that sometimes this approximation substantially underestimates the true ASN. The last study also shows that Wald's approximation of the OC-function can not be accepted in some situations.

With regard to the optimality of SPRT, Wald showed that among all sequential tests of θ_0 against θ_1 with risk levels α and β, the SPRT has the smallest possible values of $E(N;\theta_0)$ and $E(N;\theta_1)$. In practice, θ can take values other than θ_0 or θ_1, and at intermediate values, $E(N;\theta)$ can be large. Weiss (1953) showed that there are tests for which the distribution function of the decisive sampling number will never lie below the distribution function of the decisive sampling number in a SPRT. Kiefer and Weiss (1957) studied properties of such tests and gave more detailed characterizations of tests which minimize the ASN-function for a value of θ which differs from θ_0 and θ_1. This study led to the construction of tests that approximately minimize $\sup_\theta E(N;\theta)$. This can be accomplished by transforming the problem into a Bayesian decision problem. Methods have been suggested for tests of the parameter in a Bernouilli distribution (cf. Weiss (1962) and Freeman and Weiss (1964)), tests of the mean in a normal distribution with known variance (cf. Weiss (1962)) and tests of the

parameter in the exponential distribution (cf. Hoel and Mazumdar (1969)). Complex numerical calculations have to be performed, however, in order to find the two sequences $\{a_n\}$ and $\{b_n\}$ that give the boundaries to the continuation region. Anderson (1960) also used the criteria to minimize $\sup_\theta E(N;\theta)$, when making a modification of the SPRT of the mean in a normal distribution with known variance. All of these tests for minimizing $\sup_\theta E(N;\theta)$ are truncated tests, i.e. as long as N is smaller than the truncation point N_0, there is a choice between the three decisions

i) accept H_0

ii) continue sampling

iii) accept H_1.

When N equals N_0, the sampling is stopped and we choose between H_0 and H_1. The use of truncated tests also has another aspect. The decisive sampling number in a SPRT has a distribution of positive skewness. Wald (1947, p. 61) discussed the use of truncated tests in order to reduce this skewness.

I.4 Derivation and Comparison of Tests for the One-sample Case

In order to derive characteristics of truncated tests, an expression of the distribution of the decisive sampling number, N, in a SPRT has to be found. Wald (1945) derived an approximation to the d.f. of N for the case where the log likelihood ratios Z are normally distributed (see Appendices A.5 and A.6). His results can be described as follows. Let $\varphi(h)$ be the moment-generating function of Z, i.e.

$$\varphi(h) = E(e^{hZ}).$$

Now, consider the equation

$$\ln \varphi(h) = -it \qquad \text{(I.4.1)}$$

where t is a real valued parameter.

The fundamental lemma is: let Z satisfy the conditions

i) the expected value $E(Z)$ exists and differs from zero.

ii) there exists a positive number, δ, such that $P(e^Z < 1-\delta) > 0$ and $P(e^Z > 1+\delta) > 0$.

iii) for any real value of h, the expected value $E(e^{hZ}) = \varphi(h)$ exists.

Then, there exists one and only one real value, $h_0 \neq 0$, such that $\ln \varphi(h) = 0$. Furthermore $\varphi'(0)$ and $\varphi''(0)$ are both unequal to 0.

If $\varphi(h)$ is non-singular at $h = 0$ and $h = h_0$, eq. (I.4.1) has two roots, $h_1(it)$ and $h_2(it)$ for sufficiently small values of $|t|$, such that $\lim_{t\to 0} h_1(it) = 0$ and $\lim_{t\to 0} h_2(it) = h_0$.

Now define a random variable as follows: Let $\bar{S}_n = a$ if $S_n \geq a$ and $\bar{S}_n = b$ if $S_n \leq b$. Denote the difference $\bar{S}_n - S_n$ by ε. This difference ε gives the excess of S_n over a and b at termination of the test. Neglecting the excess of S_n over a and b at termination of the test means that we neglect ε, that is replace it by zero.

An approximation to the characteristic function of N, $\psi(t)$, can now be derived from the fundamental identity under the assumption of small excess of S_n over a and b. Substitute the two roots $h_1 = h_1(it)$ and $h_2 = h_2(it)$ successively in the fundamental identity, given by (cf. Wald, Appendix A.2.2)

$$E\{e^{S_n h}[\varphi(h)]^{-N}\} = 1 . \qquad \text{(I.4.2)}$$

We then get

$$\begin{aligned} Q(\theta)\, E\{e^{h_1 S_n} e^{itN} | S_n \leq b;\theta\} + [1-Q(\theta)]\, E\{e^{h_1 S_n} e^{itN} | S_n \geq a;\theta\} &= 1 \\ Q(\theta)\, E\{e^{h_2 S_n} e^{itN} | S_n \leq b;\theta\} + [1-Q(\theta)]\, E\{e^{h_2 S_n} e^{itN} | S_n \geq a;\theta\} &= 1 . \end{aligned} \qquad \text{(I.4.3)}$$

For the characteristic function of N, $\psi(t)$, we have that

$$\psi(t) = E(e^{itN}) = Q(\theta) \cdot E(e^{itN} | S_n \leq b;\theta) + [1-Q(\theta)]\, E(e^{itN} | S_n \geq a;\theta) . \qquad \text{(I.4.4)}$$

Neglecting the excess of S_n over a and b, that is replacing S_n by a and b in (I.4.3) we get the following expression of $\psi(t)$:

$$\psi(t) = E(e^{itN}) \approx \frac{e^{ah_2(it)} - e^{ah_1(it)} + e^{bh_1(it)} - e^{bh_2(it)}}{e^{ah_2(it)+bh_1(it)} - e^{ah_1(it)+bh_2(it)}} \quad . \tag{I.4.5}$$

When the log likelihood ratios Z are normally distributed with mean μ and standard deviation σ, the roots to eq. (I.4.1) can be written as

$$h_1 = d(1 - \sqrt{1 - it}) \ , \quad h_2 = d(1 + \sqrt{1 - it}) \tag{I.4.6}$$

where $d = -\mu/\sigma^2$ and the square root is determined so that the real part of $\sqrt{1 - it}$ is positive. Let

$$M = \frac{\mu^2}{2\sigma^2} N \ ; \ \xi = \sqrt{1 - it} \ .$$

The characteristic function of M can then be written as

$$\psi(t) \approx \frac{e^{ad}(e^{-bd\xi} - e^{(2ad - bd)\xi}) + e^{bd}(e^{ad\xi} - e^{(ad - 2bd)\xi})}{e^{ad}e^{bd}(1 - e^{2(ad - bd)\xi})} \ . \tag{I.4.7}$$

If $\mu > 0$, $ad < 0$ and $bd > 0$. Since the real part of $+\sqrt{1 - it}$ is greater than or equal to 1, $|e^{-2d(a - b)\xi}| < 1$ for any real value of t. The characteristic function of M can then be expanded in the form of an infinite series

$$\psi(t) = \sum_{i=1}^{\infty} r_i e^{-\lambda_i \xi} \tag{I.4.8}$$

where r_i and λ_i are constants and $\lambda_i > 0$. Each term of this series is a characteristic function except for a proportionality factor and the right-hand member of (I.4.8) can be integrated term by term so that an approximation to the distribution of M can be obtained. This approximation is a continous distribution, which depends on the fact that we have neglected the excess of S_N over a and b. The approximation will be valid when $|\mu|$ and σ are small.

When the original observations X_i follow a Poisson distribution, the moment-generating function of the likelihood ratio Z is given by

$$\varphi_{\lambda}(h) = \exp\{\lambda\,(^{\lambda_1}/_{\lambda_0})^{h} - \lambda - h\,(\lambda_1 - \lambda_0)\} \qquad (I.4.9)$$

and it is not possible to find exact analytic expressions of the two roots to (I.4.1). Different approximate analytic expression of these two roots will be studied in Section II.2.

These approximations mean that the distribution of the log likelihood ratios Z is approximated by means of a normal distribution. This in turn implies that Wald's approximation to the distribution function of N could be used.

The extent to which this approximation is good, for our purpose of deriving truncated tests, will be studied in Section IV.4. Dvoretzky et al. (1953) gave an exact expression of the moment-generating function of T — the decisive sampling time in the SPRT for testing whether a Poisson process has a parameter $\lambda = \lambda_0$ against the alternative $\lambda = \lambda_1$. This expression is too complicated to be used for deriving the distribution of T, but the three first central moments of T have been derived from it. A comparison of the moments of T and N makes it possible to study the error effected by the approximation of the roots h_1 and h_2 and the approximations of the distribution of the log likelihood ratio Z with a normal distribution. As the calculation of the moments of T required high numerical precision, comparisons were only made for $\lambda_1 = 1.25\,\lambda_0$ and $\lambda_1 = 1.5\,\lambda_0$. It is then found that the approximation used in this study gives accurate values of the first three central moments of T.

The effect of neglecting the excess of S_N is studied by comparing the four first central moments of N calculated from our approximation, with moments estimated from simulations of SPRTs. It was found that the average sampling number, $E(N,\lambda)$, was underestimated in tests on observations from a Poisson distribution. The error increased along with λ_0 and the ratio $^{\lambda_1}/_{\lambda_0}$. For the coefficient of variation, the error seemed to be small for all values studied. The approximation seems to overestimate the coefficients of skewness and excess somewhat. There was no indication of any dependence on λ_0 and λ_1 for the coefficients of variation, skewness or excess.

Expressions of the OC and ASN-function for one-sided and two-sided truncated SPRT are given in Chapter III. Simulations of such tests when $\alpha = \beta = 0.05$, $\lambda_0 = 5.0$ and 10.0 and $\lambda_1 = 1.1\,\lambda_0$, $1.25\,\lambda_0$ and $1.50\,\lambda_0$ are presented in Chapters IV and V. These simulations show that for all tests — including SPRT — the actual risk levels are lower than or equal to the intended risk levels. The selection of the truncation point in the two-sided SPRT affects the actual risk levels in the sense that a higher truncation point gives lower risks. The approximation always leads to an underestimate of the ASN-function, that increases along with λ_0 and the ratio λ_1/λ_0.

On the basis of these results different test procedures are compared in Chapter V. The ASN-function of the SPRT is lower than the size of the optimum fixed-sample test. However, the decisive sampling number in the SPRT has high variance — especially around $(\lambda_1 - \lambda_0)/2$ — and a positive coefficient of skewness.

Use of a one-sided truncated SPRT is an alternative when it is nearly certain that the values of λ are smaller than λ_0 or larger than λ_1. This test then has a smaller average sampling number than the SPRT, about the same variance and a smaller coefficient of skewness. Two-sided truncated SPRT will not reduce the maximum of the ASN-function to compensate for higher values of this function at λ_0 and λ_1. The important gains are a reduction in the variance, especially the maximum variance, and a reduction in the coefficient of skewness. The selection of the truncation point will have important effects on both of these parameters, especially near the point $(\lambda_1+\lambda_0)/2$. If the truncation point is chosen equal to 1.25 times the fixed-sample size, the decisive sampling number will have nearly the same variance for λ_0, $(\lambda_1+\lambda_0)/2$ and λ_1.

I.5 Comparing the Parameters of Two Homogeneous Poisson Distributions

Suppose that we have observations from two homogeneous Poisson distributions with parameters λ_x and λ_y. As mentioned in Section I.2, two types of models have to be used in the sequential analysis. For one type of model the observations can be analyzed pairwise, while the other type of model contains a fixed-sample as well as a sample where the observations are studied sequentially. Both of these situations are studied in Section III.3. Expressions of the characteristics of one-sided and two-sided truncated SPRT for testing

$$H_0: \lambda_y/\lambda_x \leq \delta_0 \quad \text{against} \quad H_1: \lambda_y/\lambda_x \geq \delta_1$$

are given in Section III.3.1. As in the one-sample case, only approximate expressions of the roots to the equation $\ln \varphi(h) = -it$ can be found.

The situation where the observations are not paired introduces the use of a fixed sample to estimate the parameter value of the null hypotheses. On the basis of the fixed sample, we obtain an estimate $\bar{y}$ of λ_y for a given sample size. Now define the random variable ϵ from

$$P\{|\bar{y}-\lambda_y| < \epsilon \lambda_y\} = p. \tag{I.5.1}$$

The maximal relative error in risk levels and the average sampling number, when $\lambda_x = \lambda_y$ or $\lambda_x = (1+\Delta)\lambda_y$, were calculated for different values of ϵ (cf. Section III.3.2). These calculations show that ϵ has to be small in order to avoid large changes in the risk levels. Chosing ϵ as low as 0.01 will only guarantee that the change in the risk level will be smaller than 10 per cent with probability p when $\Delta = 0.5$ and λ_y is greater than or equal to 10.0. In order to control the change in the risk levels with reasonable probability, the fixed sample has to be large. To ensure, with probability $p = 0.95$, that $\bar{y}$ will differ from λ_y by not more than $0.01\,\lambda_y$, the number of observations has to be $38,500/\lambda$, which means nearly 2,600 observations even when λ_y equals 15. Table III.3.1 shows that the number of observations needed to ensure that the

maximum relative error in the risk levels will be smaller than 0.10 is always larger than 1,000.

I.6 Heterogeneity in Parameter Values

The assumption of constant intensity for all observations in an experiment is much too simple. Several models for the distribution of road traffic accidents which take variations in the accident intensity into account were mentioned in Section I.2. One possibility concerns models which do not use other variables directly. The least restrictive of these is the multiplicative Poisson model where observations are only assumed to follow Poisson distributions, although they can have different intensities. The assumption is that observations are made when the treatment is applied and when the treatment is not applied, and that these observations can be matched pairwise. We then test whether the treatment has a constant porpotional effect in all paris of observations. Technically the use of sequential tests in this type of model involves arranging all "events" in a sequence in relation to the order in which the observations are made. Even if this cannot be accomplished, sequential tests could be used. However, the possibility of large excesses for the cumulative likelihood ratio over a and b at termination can lead to large errors in the expressions of characteristics of one-sided and two-sided truncated SPRT given in Section III.5.

Another way of allowing for variation in intensity between observations is to assume that the observations, X, are distributed according to a compound Poisson distribution. The intensity λ is then regarded as a random variable. The simplest case is when λ can be assumed to follow a gamma distribution. The observation will then follow a negative binomial distribution. This situation is studied in Section III.4.

If the intensities could be related to some dependent variable (s), regression models for the observations X could be used. Ordinary regression analysis is

based on the assumption that the observations are normally distributed. Analytic expressions of the estimates of the coefficients could not be given. Erlander, Gustavsson and Svensson (1972) proposed methods for finding numerical estimates and asymptotic simultaneous confidence regions. Comparisons of Poisson and normal regression models in the analysis of road traffic accidents, indicate that the models yield very similar results (see Erlander, Gustavsson and Lårusson (1968), Gustavsson (1968), Erlander (1970) and Gustavsson and Svensson (1972)). A simulation study made by Gustavsson and Svensson (1972) for low intensities does not clearly demonstrate any substantial advantage in using Poisson distributions instead of a normal distribution. Another aspect is that when the intensities are low, the method of grouping the observation units into homogeneous groups can be used as a good approximation to linear regression models. When the intensities are higher, ordinary regression analysis with normal distribution assumptions could be used, for which tests are presented in Gosh (1970). Use of the technique of grouping the observations units is described in Section VII.3 for the analysis of road traffic accidents (with personal injury) between motor vehicles and bicycles or mopeds in urban areas during the change-over to right-hand traffic in Sweden in 1967.

In order to study the effect of the heterogeneity of intensities in a set of observations, calculations of the OC and ASN-functions were made for various cases where the observations were assumed to follow negative binomial distributions. These were compared with OC and ASN-functions for a Poisson distribution with parameter value λ equal to the mean intensity in the set. The negative binomial distribution can be regarded as a compound Poisson distribution where λ follows a gamma distribution with mean equal to the mean intensity $\bar{\lambda}$ and variance in this case equal to $\bar{\lambda}^2/\delta$. When δ tends toward infinity, the negative binomial distribution tends toward a Poisson distribution with parameter $\bar{\lambda}$. The calculations show that the OC-function is only slightly affected even for small values of δ, but that the ASN-function is more substantially increased

when δ becomes small. The relative increase increases with λ_0, but at the same time, the number of observations necessary to separate a negative binomial distribution with δ less than or equal to δ_0 from a Poisson distribution decreases. The use of truncated tests significantly reduces the relative increases for a given value of δ as compared to the optimal fixed-sample test and the SPRT. An important advantage of using sequential tests instead of fixed-sample size tests is that when the optimum fixed-sample size is used for tests of the Poisson distribution when the true distribution is negatively binomial, the risk levels will increase substantially, even when δ is relatively large. Since logical reasons as to which model is the most appropriate are often lacking, an exploration of observations is called for. If we denote observations by x_i; $i = 1 \ldots n$, the following procedure can be used to decide on a reasonable model. In order to test whether the observations come from a Poisson distribution with constant intensity λ against the alternative of heterogeneity of intensities, study the variable

$$V = \frac{s_x^2 - \bar{x}}{\bar{x}} \qquad \text{(I.6.1)}$$

which measures how much the variance in the observations differs from the variance in a simple Poisson distribution. If s_x^2 is larger than $\bar{x}$, a negative binomial distribution with parameters μ and δ — described in Section III.4 — can be tried.

Johnson and Kotz (1969) describe different methods for estimating the parameters of a negative binomial distribution and propose two useful methods. The method of moments is feasible when δ is large but $\delta/\mu = \bar{\lambda}$ is not large and gives the estimates

$$\hat{\mu} = 1/\hat{P}\,, \quad \hat{P} = s_x^2/\bar{x} - 1 \quad \text{and} \quad \hat{\delta} = \bar{x}^2/(s_x^2 - \bar{x})\,. \qquad \text{(I.6.2)}$$

The method obtained by equating the observed and the expected number of zeros is ueful when neither δ nor $\delta/\mu = \bar{\lambda}$ is large and gives the estimates

$$\hat{\mu} = 1/\hat{P}\,, \quad \hat{\delta} = \bar{x}/\hat{P}\,, \tag{I.6.3}$$

where $\hat{P}$ is the solution to the equation

$$\frac{\beta}{\ln(1+\hat{P})} = -\frac{\bar{x}}{\ln f_0} \tag{I.6.4}$$

and f_0 is the observed proportion of zeros. These results are presented in Anscombe (1950), where tests of the departure from the negative binomial form of distribution are also given.

If $\hat{\delta}$ is large (see Chapter VI), the assumption that the observations follow a Poisson distribution and all have the same intensity will not be rejected. If $\hat{\delta}$ is small, the observations can be divided into groups which can be used in the analysis and then each group can be studied. If the observations cannot be subdivided further into groups, then other models have to be tried, e.g. the negative binomial distribution or the multiplicative Poisson model

If s_x^2 is smaller than $\bar{x}$, the absolute value of V and the test described below can be studied in order to see whether the assumption of the same intensity for all observations has to be rejected. If so, a multiplicative Poisson model could be tried.

In non-experimental applications, e.g. studies of traffic accidents, variations in intensity can follow some pattern, e.g. dependence on time. Then it is important to study the data in order to detect such patterns. This can be accomplished to some extent be testing whether the observations come from a Poisson distribution with constant intensity λ against the alternative of heterogeneity of the intensities. In addition, the relation between the observations and other variables such as time, has to be studied. A test that can be used is the dispersion test, given by the test statistic

$$D = \frac{\Sigma(x_i - \bar{x})^2}{\bar{x}} \tag{I.6.5}$$

or, if n or $(\Sigma x_i)/n$ is small, the likelihood radio test with test statistic $-2L_H$ can be used, where L_H is given by

$$L_H = -\Sigma x_i \ln x_i - (\Sigma x_i) \ln n + (\Sigma x_i) \ln (\Sigma x_i) . \qquad (I.6.6)$$

Rao and Chakravarti (1956) showed that both test statistics are approximately χ^2-distributed with (n - 1) degress of freedom and are asymptotically equivalent, but that the latter is more sensitive when n or $(\Sigma x_i)/n$ is small. The approximation to the χ^2-distribution is good when $(\Sigma x_i)/n$ is larger than 3, but can be misleading when $(\Sigma x_i)/n$ is less than unity. Critical values at 5 per cent significance levels are given for both statistics in Rao and Chakravarti's study using the exact distribution for small values of n and $(\Sigma x_i)/n$.

Chapter II APPROXIMATIONS OF THE DISTRIBUTION OF THE SAMPLE SIZE FOR A SEQUENTIAL PROBABILITY RATIO TEST OF THE POISSON MEAN

II.1 Introduction

Let X_i be a sequence of mutually independent and identically Poisson distributed random variables with the common probability mass function (p.m.f.).

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for } x = 0,1,2,3 \quad \text{(II.1.1)}$$

let S_n and the random variables Z_i be defined by

$$S_n = \sum_1^n Z_i \ , \ Z_i = \ln f(X_i;\lambda_1) - \ln f(X_i;\lambda_0) \ . \quad \text{(II.1.2)}$$

Let the sequential probability ratio test (SPRT) of the parametric hypotheses $H_0: \lambda \leq \lambda_0$ against $H_1: \lambda \leq \lambda_1$ have the operating characteristic function $Q(\lambda)$. This test operates as follows. Continue sampling as long as

$$b < S_n < a \ .$$

Stop sampling and accept H_0 as soon as $S_n \leq b$ and stop sampling and accept H_1 as soon as $S_n \geq a$. The constants a and b should be chosen in such a way that $Q(\lambda_0) \geq 1-\alpha$ and $Q(\lambda_1) \leq \beta$. Denote the number of observations necessary to terminate the test — the decisive sampling number — as N, which will be a random variable. When the random variables Z_i are studied below, the index is dropped and the notation Z is used.

The purpose of this chapter is to obtain expressions of the cumulative distribution function of N, which could be used to derive truncated tests.

Wald (1947) gives an approximate formula for the characteristic function of N (cf. Section I.4). His derivation neglects the excess of S_N over the boundaries a and b. The error due to this neglect will be small if $|E(Z)|$ and $E(Z^2)$ are small. If Z is normally distributed, Wald also gives an approximation of the distribution of N. In passing, Wald points out that this approximation could be used even if Z is not normally distributed, given that the mean and the standard deviation of Z are small as compared to a and b. His argument rests on the central limit theorem, which could be used if we define new random variables

$Z_j^* = \sum_{i=(j-1)r+1}^{jr} Z_i$ $(i = 1,2\ldots)$. If the Z_i's are independent and identically distributed, Z_j^* will be nearly normally distributed for large r.

In the Poisson case this means that we construct new random variables $X_j^* = \sum_{(j-1)r+1}^{jr} X_i$ with the parameter $\lambda^* = r \cdot \lambda$ which will be approximately normally distributed if λ^* is large enough. Of course this will reduce the advantage of the sequential test, as the test can be terminated only after we have taken $r, 2r, 3r\ldots$ of the original observations.

Gosh (1970) proposes that the method used by Wald for the normal distribution could be applied for other distributions. According to Section I.4 the problem then is to obtain analytic expressions of the roots to the equation

$$\ln \varphi(h) = -it \qquad \text{(II.1.3)}$$

where $\varphi(h) = E(e^{hZ})$ is the momentgenerating function of Z. When Z is normally distributed this equation is a second-degree polynomial in h. In other cases this equation is not a second-degree polynomial and the roots usually cannot be expressed as analytic functions of the parameters. Gosh does not discuss this problem. Different analytic expressions that could be used as approximations to the roots of equation (II.1.3) will be discussed in Section II.2.

Bhate (1959) derives some lower and upper bound for the cumulative distribution function of N, although not in a form that could be used to derive truncated tests analytically.

Chanda (1971) has shown that if the distribution of the original observations belongs to the so-called Darmois-Koopman exponential class, a certain function of N is asymptotically normally distributed when $|\lambda_1 - \lambda_0|$ tends toward zero. With the transformation $\theta = \ln \lambda$, the test differentiates between the two hypotheses $H_0: \theta \leq \theta_0$ and $H_1: \theta \geq \theta_1 = \theta_0 + \Delta$. It is then possible to find constants $A > 0$ and $B > 0$ such that $[N\Delta - A]/B\sqrt{\Delta}$ tends toward $N(0,1)$ as $\Delta \to 0$ when $\theta < \theta_0$ or $\theta > \theta_1$. The constant B is finite and both constants depend on

a if $\theta < \theta_0$ and on -b if $\theta > \theta_0$. Unfortunately, the error made by using this asymptotic normal distribution as an approcimation to the true distribution of N (when $\Delta > 0$) depends on the values of both θ and Δ. Elofsson (1974) has shown that if θ is written as $\theta = \theta_0 + c\Delta$, it is necessary that both $\Delta \to 0$ and $|c| \to \infty$ so that $c\Delta \to k > 0$ in order for the asymptotic result to hold. This result indicates that the asymptotic distribution cannot be used in any case when $\lambda_0 \leq \lambda \leq \lambda_1$, which is of great importance when deriving truncated tests.

The disposition of Chapter II is as follows. Different expressions of the roots to the equation $\ln \phi(h) = -it$ are discussed and studied in Section II.2. An approximate expression of the distribution of N is given in Section II.3, where the probability mass function of N is approximated with a probability density function.

II.2 Analytic Expressions of the Roots to the Equation $\ln \varphi(h) = -it$

It follows from Section II.1 that for the log likelihood ratio Z, used when testing the hypotheses $H_0: \lambda \leq \lambda_0$ against $H_1: \lambda \geq \lambda_1$ we have

$$\begin{aligned} Z &= X_i \ln (\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0) \\ E(Z) &= \lambda \cdot \ln (\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0) \\ V(Z) &= \lambda \cdot \ln^2 (\lambda_1/\lambda_0) \end{aligned} \qquad \text{(II.2.1)}$$

where X is Possion distributed with parameter λ.

The moment-generating function of Z is given by

$$\varphi_\lambda(h) = E(e^{hZ}) = \exp\{\lambda (\frac{\lambda_1}{\lambda_0})^h - \lambda - h(\lambda_1 - \lambda_0)\} . \qquad \text{(II.2.2)}$$

In order to use Gosh's (1970) method of approximating the distribution function of N, which is the same as that used by Wald when Z is normally distributed (cf. Section I.4), we have to find analytic expressions of the roots $h_1(it)$ and $h_2(it)$ to the equation

$$\ln\varphi_{\lambda}(h) = -it \ . \qquad \text{(II.2.3)}$$

On the basis of the fundamental lemma, Wald (1947, p. 186) showed that when $E(Z) \neq 0$ for sufficiently small values of $|t|$, this equation has two roots such that $\lim_{t\to 0} h_1(it) = 0$ $\lim_{t\to 0} h_2(it) = h_0$ where h_0 is the non-zero solution to the equation $\ln\varphi(h) = 0$.

If $\ln\phi(h)$ is approximated by means of a second-degree polynomial in h, this implies that the distribution of Z will be approximated by a normal distribution. We then have to find expressions of the mean and the variance of this normal distribution, which can be done in various ways. The conditions that the root h_0 should satisfy are given in (II.2.7). The approximation of the equation by $\varphi_{\lambda}(h) = 0$ by means of a second-degree polynomial in h means that h_0 will depend on the ratio between the meand and the variance of the corresponding normal distribution. If the mean of this normal distribution is fixed to E(Z) the conditions in (II.2.7) below cannot be satisfied, but this can be achieved if we let the mean differ from E(Z) while the variance equals V(Z). One reason for choosing the variance equal to V(Z) lies in the fact that if $(\lambda_1/\lambda_0)^h$ is Taylor expanded round $h = 0$ and all terms including the third and higher powers of h are neglected, an expression of type (II.2.4) below is obtained.

In the following, the variance in this distribution has been chosen equal to V(Z) and different values of the mean of this distribution have been studied in order to find an approximation of the distribution of N.

This means that equation (II.2.3) will be approximated to

$$\frac{\lambda \cdot \ln^2(\lambda_1/\lambda_0)}{2}(h^2 - h_0 h) + it = 0 \qquad \text{(II.2.4)}$$

where the roots expressed in h_0 are

$$h_{1,2}(it) = \frac{h_0}{2}\left\{1 \pm \sqrt{1 - \frac{2it \cdot 4}{h_0^2 \cdot \lambda \cdot \ln^2(\lambda_1/\lambda_0)}}\right\} \qquad \text{(II.2.5)}$$

where the sign of the square root is determined so that the real part of it is positive.

We now have to find a useful approximation of h_0. In accordance with Wald's proposal (pp. 195-196), we use E(Z) to determine h_0. This seems reasonable in any case where λ is large (cf. Section I.4).

Chanda (1971) had to approximate the moment-generating function and for the Poisson distribution this meant that he expanded $(\lambda_1/\lambda_0)^h$ around $h = 0$ and then approximated the moment-generating function of Z with a second-degree polynomial in h. Both of these approximations will give us the same expression for h_0, that is

$$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)} \quad . \qquad \text{(II.2.6)}$$

When deriving the truncation point and the continuation region of truncated tests, only the distribution function of N at $\lambda = \lambda_0$ and $\lambda = \lambda_1$ are used. It would then be important to obtain expressions of h_0 that give the correct value of $\varphi(h)$ in these points. For $\lambda = \lambda_0$ it is known that h_0 equals 1 and for $\lambda = \lambda_1$ it equals -1. We also know that h_0 is positive for $\lambda < \lambda'$ and negative for $\lambda > \lambda'$, where

$$\lambda' = \frac{\lambda_1 - \lambda_0}{\ln(\lambda_1/\lambda_0)}$$

is the value of λ where $E(Z) = 0$. If an expression of $h_0(\lambda)$ which fulfils the conditions $h_0(\lambda') = 0$ can be found, then it will be easier to control the error in the distribution function of N when $\lambda_0 < \lambda < \lambda_1$, especially when λ lies near λ'. From the expression of the characteristics of truncated SPRT, it can be seen that the boundaries of the continuation region as well as the truncation point depend only on the distribution of N when λ equals λ_0 and λ_1. For the derivation of truncated tests, it then seems less important that $h_0(\lambda') = 0$

as compared to having an expression that fullfils the conditions

$$\begin{aligned} h_0(\lambda_0) &= 1 \\ h_0(\lambda_0) &= -1 \ . \end{aligned} \qquad \text{(II.2.7)}$$

If h_0 is expressed as a second-degree polynomial in λ all three conditions can be fulfilled. Since the coefficients of this expression will not be simple, simpler expressions which only satisfy (II.2.7) have been used. For our purposes, it seems that little is gained by using a second-degree polynomial in λ. Two expressions of h_0 that satisfy (II.2.7) will then be studied and compared with expression (II.2.6). These two expressions are

$$\begin{aligned} h_0(\lambda) &= a + b\lambda \\ h_0(\lambda) &= \frac{m}{\lambda} - r \ . \end{aligned} \qquad \text{(II.2.8)}$$

Solving for the coefficients in order for these expressions to satisfy conditions (II.2.7) and using (II.2.6) we obtain three expressions of h_0

$$\begin{aligned} h_0 &= \frac{-2\left[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0\right]}{\lambda \cdot \ln^2(\lambda_1/\lambda_0)} \\ h_0 &= \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)} \\ h_0 &= \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)} \end{aligned} \qquad \text{(II.2.9)}$$

In order to compare these expressions, we begin by studying how well they approximate the correct solution to the equation $\varphi_\lambda(h) = 0$, where $\varphi_\lambda(h)$ is given in (II.2.2). This is done by calculating the value of $\varphi_\lambda(h)$ for each expression. These values cannot be compared unless we know if the values of h_0 given by (II.2.9) are all larger or smaller than the value of h, h_{min}, which gives the minimum of $\ln\varphi_\lambda(h)$. Introducing $c = (\lambda_1 - \lambda_0)/\lambda_0$ and writing λ as $\lambda_0(1 + k\Delta)$ makes it possible to study all of the expressions, including h_{min}, as functions of c and k. The values of k for which the expressions in

(II.2.9) all give values that lie on the same side of h_{min} are studied in Appendix A. It is then found that for actual values of Δ, that is $0 < \Delta < 1/2$, the results in Appendix A can be summarized in Figure II.2.1, which contains graphs of the four functions.

Figure II.2.1 shows that all of the expressions give values of h_0 which lie on the same side of h_{min} except in the area around $k = 0.5$. The size of this area depends on the value of Δ. If the expression

$$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1+\lambda_0)\lambda}{\lambda(\lambda_1-\lambda_0)}$$

is disregarded this interval becomes smaller. This makes it possible to compare the two expressions in a wider range. The values of h_0 for the three expressions in (II.2.9) and the connected values of $\ln\varphi_\lambda(h)$ for different values of λ_0, λ_1 and λ are presented in Appendix B.

The tendencies are the same for all values of λ_0 and λ_1 but the differences are most pronounced when λ_0 is large. For λ greater than $\frac{\lambda_1+\lambda_0}{2}$ it is quite clear from Figure II.2.1 and the values given in the Appendix B that the approximation

$$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1+\lambda_0)\lambda}{\lambda(\lambda_1-\lambda_0)}$$

is a poor alternative. Figure II.2 shows that this is also the case when λ is smaller than or equal to $(\lambda_1+\lambda_0)/2$. Comparing the two other approximations we see that the expression

$$h_0 = \frac{\lambda_1+\lambda_0-2\lambda}{(\lambda_1-\lambda_0)}$$

seems to give the best approximation as long as $k < 0.2\text{-}0.3$ and $k > 0.7\text{-}0.8$, where k is given by $\lambda = \lambda_0(1+k\Delta)$. For intermediate values of k, the expression

$$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0 - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$$

will give the best approximation.

In order to see how the choice of approximations of the roots influences the approximation of the distribution of N, we study the effect on the approximation to the moments of N. Table II.2.1 contains the values of the first four central moments of N using the approximate d.f. for N with h_0 given by formula (II.2.14), h_{0I}, and by (II.2.15), h_{0II}. A comparison with values calculated from an exact expression of the moment-generating function of T and values estimated from simulations of SPRT (cf. Section II.5) shows that the errors are smaller using h_{0I} than h_{0II}, for $\lambda = \lambda_0$ and λ_1, especially when λ_0 and $\lambda_1 - \lambda_0$ are large. The difference is more pronounced for higher moments. Another indication for not using h_{0II} is the fact that this expression gives erroneous relations between the moment of N for $\lambda = \lambda_0$ and $\lambda = \lambda_1$. From the exact expression of the moment-generating function of T, we find that

$$\begin{aligned} &E(T;\lambda_0) > E(T;\lambda_1) \\ &V(T;\lambda_0) < V(T;\lambda_1) \\ &[E(T-\mu_T)^3;\lambda_0] < [E(T-\mu_T)^3;\lambda_1] . \end{aligned} \tag{II.2.16}$$

Using h_{0I} the relations will be

$$\begin{aligned} &E(N;\lambda_0) > E(N;\lambda_1) \\ &V(N;\lambda_0) \simeq V(N;\lambda_1) \\ &[E(N-\mu_N)^3;\lambda_0] \simeq [E(N-\mu_N)^3;\lambda_1] . \end{aligned} \tag{II.2.17}$$

We have the following relations for h_{0II}, where the differences are more pronounced when $|E(Z;\lambda)|$ and $|V(Z;\lambda)|$ are large:

$$\begin{aligned} &E(N;\lambda_0) \gg E(N;\lambda_1) \\ &V(N;\lambda_0) > V(N;\lambda_1) \\ &[E(N-\mu_N)^3;\lambda_0) > [E(N-\mu_N)^3;\lambda_1] \end{aligned} \tag{II.2.18}$$

where $\gg$ means that the difference is larger than in (II.2.16).

As has been shown above (see Figure II.2.1) h_{0II} gives more accurate values in an intervall around $\lambda = (\lambda_1+\lambda_0)/2$. This means that an optimal rule for a truncated test would be to use h_{0I} to estimate the contination limits and the truncation point — which depends solely on $\lambda = \lambda_0$ and λ_1 — and then use h_{0II} to calculate an estimate of E(N) when λ lies near $\lambda = (\lambda_1+\lambda_0)/2$.

II.3 An Approximation to the Distribution of N

In the preceding section, we found that using the expression

$$h_0 = \frac{\lambda_1+\lambda_0-2\lambda}{(\lambda_1-\lambda_0)}$$

as an approximation to the non-zero root to the equation $\ln \varphi(h) = 0$ was the best alternative for deriving truncated tests. This approximation also implied that the distribution of Z was approximated with a normal distribution with mean $-\frac{1}{2} h_0 \cdot \lambda \cdot \ln^2(\lambda_1/\lambda_0)$ and variance $\lambda \ln^2(\lambda_1/\lambda_0)$. This implies that Wald's approximation to the distribution function of N can be used and if suitable transformation are made, the expression given by Gosh (1970, p. 217) for the approximate d.f. of N can be used when testing the mean in a normal distribution with known variance. This gives

$$p(n,\lambda) \approx \frac{e^{-n\eta^2/2}}{n\sqrt{2\pi n}} \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k \lambda_{kj} e^{-\lambda_{kj}^2/2n} \quad \text{when } \lambda = \lambda' \qquad \text{(II.3.1)}$$

where

$$\eta = \frac{(\lambda_1+\lambda_0-2\lambda)\sqrt{\lambda \cdot \ln(\lambda_1/\lambda_0)}}{2(\lambda_1-\lambda_0)}$$

$$\delta_1 = -\delta_2 = -e^{b'\eta} \qquad \delta_3 = -\delta_4 = -e^{a'\eta}$$

$$\lambda_{1j} = 2(j+1)a' - (2j+1)b' \qquad \lambda_{2j} = 2ja' - (2j+1)b' \qquad \text{(II.3.2)}$$

$$\lambda_{3j} = (2j+1)a' - 2(j+1)b' \qquad \lambda_{4j} = (2j+1)a' - 2jb' \ .$$

With a one-sided test and an upper boundary for S_N, that is $b \to -\infty$ and a finite,

$$p(n;\lambda) \approx \frac{a'}{n\sqrt{2\pi n}}\, e^{\frac{(-n\eta - a')^2}{2n}} \quad \text{for all } \lambda > \lambda' \ . \tag{II.3.3}$$

With an one-sided test and a lower boundary for S_N, that is $a \to \infty$ and b finite, an approximation to the p.m.f. of N will be

$$p(n;\lambda) = \frac{-b'}{n\sqrt{2\pi n}}\, e^{\frac{(-n\eta - b')^2}{2n}} \quad \text{when } \lambda < \lambda' \ . \tag{II.3.4}$$

Table II.2.1 Values of the first four central moments of N (the decisive sampling number) when using different expressions to the non-zero solution of the equation $\ln \varphi_\lambda(h) = 0$ to derive an approximation to the d.f. of N.

$$h_{0I} = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)} \qquad h_{0II} = \frac{2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \cdot \ln^2(\lambda_1/\lambda_0)}$$

			$E(N) = \mu_N$		$V(N)$		$E(N - \mu_N)^3$		$E(N - \mu_N)^4$	
λ_0	λ_1	λ	h_{0I}	h_{0II}	h_{0I}	h_{0II}	h_{0I}	h_{0II}	h_{0I}	h_{0II}
1.00	1.10	1.00	565.05	570.53	$1.61 \cdot 10^5$	$1.73 \cdot 10^5$	$1.22 \cdot 10^8$	$1.55 \cdot 10^8$		
		1.10	547.38	541.88	$1.61 \cdot 10^5$	$1.47 \cdot 10^5$	$1.22 \cdot 10^8$	$0.83 \cdot 10^8$		
1.00	1.25	1.00	98.67	100.85	4692	5500	$6.07 \cdot 10^5$	$9.59 \cdot 10^5$	$1.88 \cdot 10^8$	$3.98 \cdot 10^8$
		1.25	91.60	89.40	4692	3708	$6.07 \cdot 10^5$	$0.97 \cdot 10^5$	$1.88 \cdot 10^8$	
1.00	1.50	1.00	28.03	29.11	355	458	12638	24358	$1.07 \cdot 10^6$	$2.94 \cdot 10^6$
		1.50	24.49	23.39	355	207	12643		$1.07 \cdot 10^6$	
5.00	5.50	5.00	113.01	114.11	6433	6934	$0.97 \cdot 10^6$	$1.24 \cdot 10^6$	$3.52 \cdot 10^8$	$5.45 \cdot 10^8$
		5.50	109.48	108.38	6433	5888	$0.97 \cdot 10^6$	$0.66 \cdot 10^6$	$3.52 \cdot 10^8$	$1.13 \cdot 10^8$
5.00	6.25	5.00	19.73	20.17	188	221	4859	7733	$3.00 \cdot 10^5$	$6.44 \cdot 10^5$
		6.25	18.32	17.88	188	147	4859	655	$3.00 \cdot 10^5$	
5.00	7.50	5.00	5.61	5.82	14.19	18.53	101	201	1717	4881
		7.50	4.90	4.68	14.20	7.79	101		1718	
10.00	11.00	10.00	56.51	57.05	1608	1734	$1.22 \cdot 10^5$	$1.55 \cdot 10^5$	$2.20 \cdot 10^7$	$3.41 \cdot 10^7$
		11.00	54.74	54.19	1608	1471	$1.22 \cdot 10^5$	$0.83 \cdot 10^5$	$2.20 \cdot 10^7$	$0.70 \cdot 10^7$
10.00	12.50	10.00	9.87	10.08	46.92	55.32	607	976	18750	40779
		12.50	9.16	8.94	46.92	36.61	607	62	18752	
10.00	15.00	10.00	2.80	2.91	3.55	4.70	12.64	25.90	107.30	317.30
		15.00	2.45	2.34	3.55	1.79	12.64		107.35	
15.00	16.50	15.00	37.67	38.04	715	771	$3.61 \cdot 10^4$	$4.61 \cdot 10^4$	$4.35 \cdot 10^6$	$6.75 \cdot 10^6$
		16.50	36.49	36.13	715	654	$3.61 \cdot 10^4$	$2.44 \cdot 10^4$	$4.35 \cdot 10^6$	$1.36 \cdot 10^6$
15.00	18.75	15.00	6.58	6.72	20.85	24.66	180	292	3704	8165
		18.75	6.11	5.96	20.85	16.15	180	12	3704	

Figure II.1 Graph of the approximations to the non-zero root to the equation $\ln\varphi_\lambda(h) = 0$ and the value of h, h_{min}, for which the minimum of $\ln\varphi_\lambda(h)$ is reached.

$$h_B = \frac{\lambda_1+\lambda_0-2\lambda}{(\lambda_1-\lambda_0)} \qquad h_A = \frac{-2\left[\ln(\lambda_1/\lambda_0)-\lambda_1+\lambda_0\right]}{\lambda\cdot\ln^2(\lambda_1/\lambda_0)}$$

$$h_C = \frac{2\lambda_1\lambda_0-\lambda(\lambda_1+\lambda_0}{\lambda(\lambda_1-\lambda_0)}$$

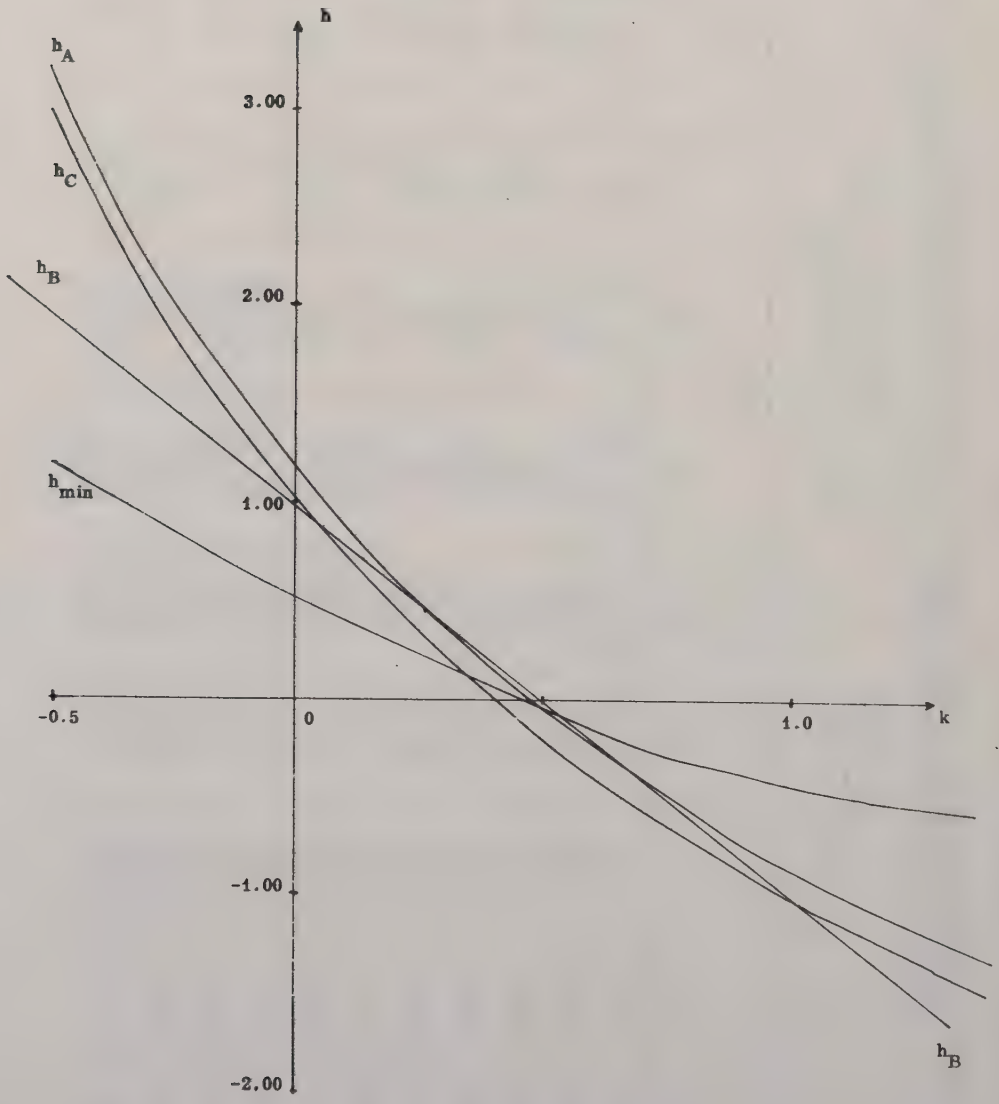
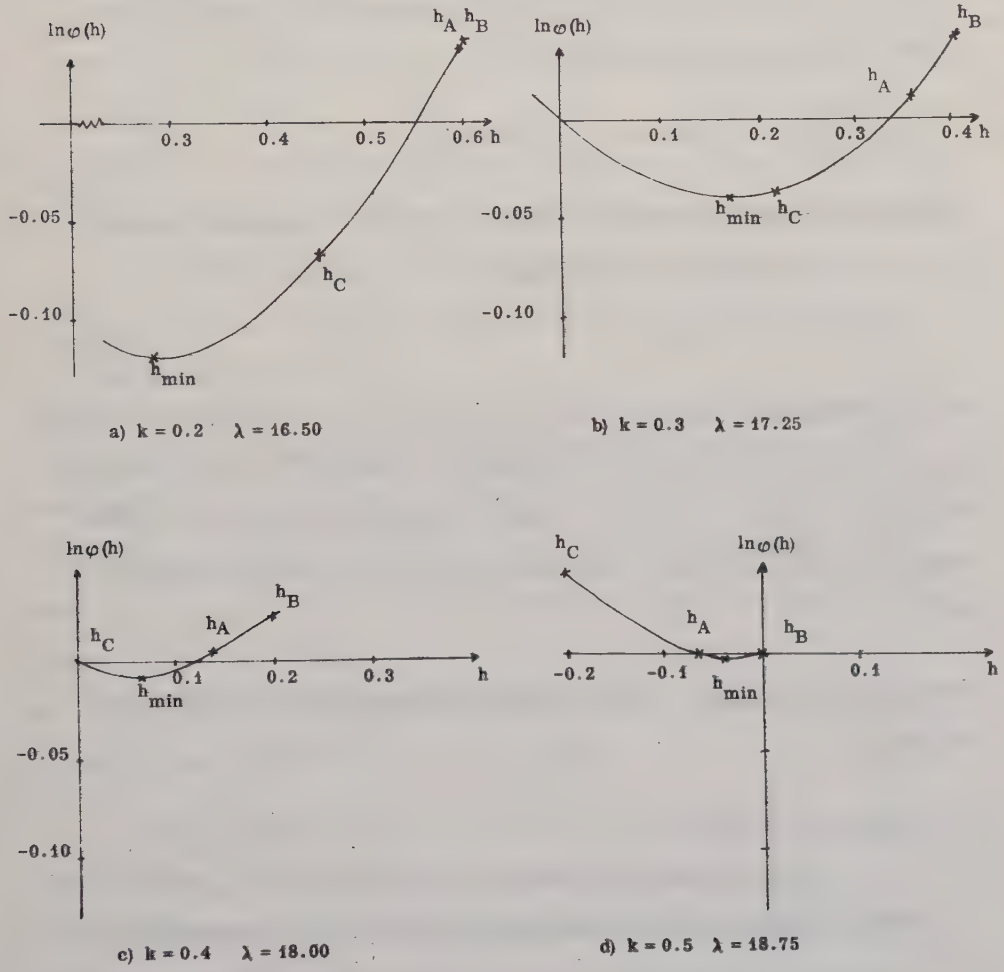


Figure II.2.2 Graphs of the function $\ln\varphi_\lambda(h)$ for some values of λ when $\lambda_0 = 15.0$ and $\lambda_1 = 22.50$. The graphs are used to show how well different expressions $h_A(\lambda)$, $h_B(\lambda)$ and $h_C(\lambda)$ approximate the non-zero root to the equation $\ln\varphi_\lambda(h) = 0$ round the value $\lambda = (\lambda_1+\lambda_0)/2$. The expressions are defined in Figure II.2.1



Chapter III Truncated Sequential Tests

III.1 Generalized SPRT and Other Optimal Test Procedures

According to Gosh (1970) a generalized SPRT is defined by the fact that the region of sampling continuation is bounded by two lines (which do not necessarily have to be parallel to the n-axes) through b and a and a truncation point $N = N_0$. We thus obtain a class C of sequential test procedures for testing H_0 against H_1.

As before, let S_n and Z_i be defined by

$$S_n = \sum_1^n Z_i = \Sigma \{\ln f(X_i, \lambda_1) - \ln f(X_i, \lambda_0)\} .$$

A member of class C, denoted $S(b, a, c_1, c_2, N_0, c)$ is defined as follows:

For $n < N_0$ continue sampling as long as

$$b + c_1 n < S_n < a + c_2 n . \qquad \text{(III.1.1)}$$

Stop sampling and accept H_0 as soon as $S_n \leq b + c_1 n$; stop sampling and accept H_1 as soon as $S_n \geq a + c_2 n$. If no decision is reached for $n < N_0$, stop sampling when $n = N_0$, if $S_{No} < c$ accept H_0, otherwise accept H_1. The six constants should be chosen in such a way that the test has the properties in (III.1.3).

Gosh divides the class C of "restricted sequential procedures" into five subclasses:

C_1 consists of all tests with $N_0 < \infty$, $b + c_1 n = a + c_2 n = \infty$ for all $n < N_0$, which means that a decision is only made after N_0 observations have been made. This subclass then contains all fixed-sample tests.

C_2 consists of all tests with $N_0 = \infty$, $c_1 = c_2 = 0$, which means that the test is not truncated and the continuation region is bounded by two lines through a and b parallel to the n-axis. This subclass contains sequential probability ratio tests.

C_3 consists of all tests with $N_0 < \infty$, $b + c_1 n = -\infty$ and $-\infty < a + c_2 n < \infty$ for all n, which means that the continuation region is bounded by an upper

boundary and the test is truncated. This subclass contains one-sided sequential tests where H_1 can be accepted for any $N = 1, \ldots, N_0$, but H_0 can only be accepted for $N = N_0$.

C_4 consists of all tests with $N_0 < \infty$, $-\infty < b + c_1 n < \infty$ and $a + c_2 n = \infty$ for all n, which means that the continuation region is bounded by a lower boundary and the test is truncated. This subclass contains one-sided tests, where H_0 can be accepted for any $N = 1, \ldots, N_0$, but H_1 can only be accepted for $N = N_0$.

C_5 consists of all test in $C\text{-}C_1\text{-}C_2\text{-}C_3\text{-}C_4$, which means nontruncated tests with two finite lines as boundaries to the contination region, both of which are not parallel to the n-axis and truncated tests with two finite lines as boundaries to the continuation region. Among other tests this subclass contains two-sided SPRT.

In order to choose an optimal sequential test procedure, the risk levels for the test should be fixed an the optimality of the test is related to the distribution of N, the decisive sampling number (cf. Section I.3). Then assume that observations X_i $(i = 1, 2 \ldots)$ are independent and identically distributed with probability function $f(x;\theta)$. The problem is to choose between hypotheses $H_0: \theta \leq \theta_0$ and $H_1: \theta \geq \theta_1$. To construct SPRTs for these composite hypotheses, it has to be assumed that the log likelihood ratio, given by

$$Z = \ln \left\{ \frac{f(x, \theta_1)}{f(x, \theta_0)} \right\} \qquad \text{(III.1.2)}$$

is a monotonic function of θ. We then study test procedures having the properties

$$\begin{aligned} &P(\text{accept } H_0 \mid \theta) \geq 1 - \alpha \quad &&\text{when } \lambda \leq \lambda_0 \\ &P(\text{accept } H \mid \theta) \leq \beta \quad &&\text{when } \lambda \geq \lambda_1 . \end{aligned} \qquad \text{(III.1.3)}$$

The OC-function is given by

$$Q(\theta) = P(\text{acc } H_0 | \theta)$$

and must satisfy condition (III.2.3).

From among these procedures, we have to choose the one which is optimal with respect to the distribution of N in some sense.

It then seems accurate to concentrate on the mean of N. As a function of θ, this mean is called the ASN-function (ASN = average sampling number). One property of Wald's "sequential probability ratio test" (SRPR) is that it has the smallest value of E(N) when $\theta = \theta_0$ and $\theta = \theta_1$. If the true value of θ lies between θ_0 and θ_1, E(N) can, however, take considerably high values. From among tests which have the property in (III.1.2), several authors (e.g. Anderson (1969) and Donnerly (1957)) have choosen the one that gives the lowest maximum of the ASN-function. Anderson studied tests of the mean μ in a normal distribution with known variance and derived characteristics of truncated tests belonging to class C. Weiss (1953) studied a more general class of tests, where the boundaries to the continuation region did not have to be lines. Kiefer and Weiss (1957) have shown that it is possible to find an approximate optimal test having the property that it gives the lowest maximum of the ASN-function. They demonstrated that this was a truncated test and could be found using the following method: First, determine exactly or approximately the value λ where the ASN-curve of SPRT has its maximum. Then choose the test that minimizes the value of E(N) at this point. The problem of finding a test with the properties in (III.1.3) and minimizing the value of the ASN-curve at a point λ_2 where $\lambda_0 < \lambda_2 < \lambda_1$ can be transformed into the following decision problem. Let a_0 be the decision that H_0 is accepted and a_1 the decision that H_1 is accepted.

Let $L(\lambda, a_i)$ $(i = 1; 2)$ be the loss using strategy a_i when λ is the true parameter value. Let $C(\lambda)$ denote the cost of taking one observation. The loss function is

$$L(\lambda, a_0) = \begin{cases} 1 & \text{when } \lambda = \lambda_1 \\ 0 & \text{otherwise} \end{cases}$$

(III.1.4)

$$L(\lambda, a_1) = \begin{cases} 1 & \text{when } \lambda = \lambda_0 \\ 0 & \text{otherwise.} \end{cases}$$

The observation cost is

$$C(\lambda) = \begin{cases} 1 & \text{if } \lambda = \lambda_2 \\ 0 & \text{otherwise.} \end{cases} \qquad \text{(III.1.5)}$$

Given an a priori distribution with weights b_0, b_1 and b_2 to λ_0, λ_1 and λ_2 we have a one-to-one correspondence between these weights and the risks α and β.

The Bayes' strategy of this decision problem is a test having the property (III.1.1) and minimizing the ASN-curve at $\lambda = \lambda_2$. Hoel and Mazumdar (1969) have shown how a test can be found which approximately minimizes the maximum of the ASN-curve when the X_i's are exponentially distributed.

However, to my knowledge no analytic expression of the bounds to the continuation region, truncation point and the OC and ASN-functions have been given. The computation procedures can be simplified by using an invariance property of the Bayes' stopping region, although the computations would still be extensive.

It is often unrealistic to study only the ASN-curve of a test. For the SPRT, it can be shown that N has often large variance and positive skewness, that is $P(N < N_0)$ does not necessarily lie near one even when N_0 is much larger than E(N). Gosh (1970), pp. 142-4) assigned an upper bound to this probability. However, in many applications it is not realistic to let the sample size N vary too much. Truncated tests have to be used instead. This involves taking observations sequentially but stopping after N_0 observations if we have not decided between the two alternatives before N_0 observations have been made. When the

upper bound N_0 is high (cf. Gosh (1970)), the OC-curve and ASN-curve will not change very much. But if we want a lower truncation bound, it is not longer possible to approximate the OC and ASN-curves in the usual way.

The difficulty in finding an optimal criterion which can always be used motivates a study of different types of sequential test procedures. Gosh (1970) made this kind of study for the case where the X_i's are normally distributed, with the respect to all "restricted sequential procedures" (see Armitage (1957)). He also made numerical comparisons between tests that are "best" or "nearly" best from different points of view.

III.2 Truncated SPRT of Poisson Parameter

Given observations X_i $(i = 1,2\ldots)$ that are independent and identically distributed according to a Poisson distribution

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \text{for } x = 0,1,\ldots$$

we want to choose between the hypotheses $H_0:\lambda \leq \lambda_0$ and $H_1:\lambda \geq \lambda_0$ with given risk levels α and β. An approximate distribution of N was derived in Chapter II using results from Wald (1947). An approximate analytic expression of the non-zero root h_0 of the equation $\ln \phi_\lambda(h) = 0$ had to be used to find this approximation. The distribution of N at $\lambda = \lambda_0$ and $\lambda = \lambda_1$ will be used to determine truncation point and boundaries to the continuation region for truncated tests. The analytic expression of h_0 is then chosen in such a way that it gives the correct values for $\lambda = \lambda_0$ and λ_1.

Wald's derivation of the distribution of N was performed under the assumption that the log likelihood ratios, Z, are normally distributed. For the Poisson case, the distribution of the log likelihood ratios was approximated using a normal distribution. This made it possible to use Wald's result, which was given more explicitly later on by Gosh (1970). Wald's derivation rests on the assumption that the excess of $S_n = \sum_1^n z_i$ at the termination of the test is neglected

(cf. Section I.4). This assumption also adheres to the result given for truncated tests and will lead to underestimates of the moments of N. The size of this underestimate depends on the size of $|E(Z)|$ and $V(Z)$ in relation to the boundaries a and b of the continuation region and will be studied in Section IV.3.

We now turn to tests belonging to the subclasses C_3, C_4 and C_5. The argument set forth by Gosh (1970, pp. 233-34, 236) can be used to decide which tests should be studied. The two subclasses C_3 and C_4 are symmetrical and Gosh concentrates on one of them, C_3, according to the following line of reasoning (p. 234):

If (a, c_2, N_0) satisfies $Q(\lambda_0) = 1-\alpha$ and $Q(\lambda_1) = \beta$, the resulting member of C_3 is admissible for testing H_0 against H_1. As long as $\alpha + \beta < 1$, more than one combination of the three parameters will satisfy the two equations. It can be shown directly from the expression of $E(N;\lambda)$ that for any fixed value of one of the three constants, the test in C_3 that satisfies the strict equalities in (III.2.1) is more efficient than the test that satiesfies at least one of the two strict inequalities in (III.2.1)

$$Q(\lambda_0) \geq 1-\alpha \qquad Q(\lambda_1) \leq \beta \,. \tag{III.2.1}$$

Consequently, in order to determine the best member of C_3, we need only consider the subclass $C_3' \subset C_3$ characterized by

$$Q(\lambda_0) = 1-\alpha \quad \text{and} \quad Q(\lambda_1) = \beta \,.$$

The member of C_3' which can be found most simply is the test $S(a, 0, N_0)$ where $a = \ln \frac{1-\beta}{\alpha}$ and N_0 is the solution to the equation

$$\phi\left(\frac{a}{\delta\sqrt{N_0}} - \frac{\delta\sqrt{N_0}}{2}\right) - \frac{(1-\beta)}{\alpha}\,\phi\left(-\frac{a}{\delta\sqrt{N_0}} - \frac{\delta\sqrt{N_0}}{2}\right) = \beta \tag{III.2.2}$$

Here δ is a function of $\lambda, \lambda_0, \lambda_1$.

The left-hand side of (III.2.2) as a function of $\delta\sqrt{N_0}$ is strictly decreasing fron one to zero and therefore (III.2.2) has a unique solution for every $0 < \beta < 1-\alpha$. That the test $S(a, 0, N_0)$ defined in this way belongs to C_3' can be seen from the following argument.

If H_1 is accepted for some N, such that $0 < N \leq N_0$ we have that $S_N = a = \ln \frac{(1-\beta)}{\alpha}$. There is a theorem which says that for every closed and open test the OC-($Q(\lambda)$) and power-($P(\lambda)$) functions satisfy

$$E(e^{S_N} | H_0 \text{ accepted}) = \frac{Q(\lambda_1)}{Q(\lambda_0)} \qquad \text{(III.2.3)}$$

$$E(e^{S_N} | H_1 \text{ accepted}) = \frac{P(\lambda_1)}{P(\lambda_0)} = \frac{1-Q(\lambda_1)}{1-Q(\lambda_0)} . \qquad \text{(III.2.4)}$$

As we have $E(e^{S_N} | H_1 \text{ accepted}) = \frac{(1-\beta)}{\alpha}$ and $Q(\lambda_1) = \beta$, from (III.2.4) we get $Q(\lambda_0) = 1-\alpha$, that is the test belongs to C_3'.

This test will <u>not</u> be the best member in C_3' in the sense that there exist tests in C_3' which require a smaller average sampling number, at least for some values of λ. However, Gosh conjectures that a UME test within C_3' does not exist.

According to Gosh it seems reasonable to fix N_0 beforehand and then select a and c_2 which give a test that achieves the risks (α, β) and minimizes $E(N; \lambda)$ for some specified λ. This optimization problem is difficult and hardly worthwhile since the subclass C_3 (or C_4) does not have any known optimum property vis-à-vis other subclasses.

For tests in subclass C_5, six constants should be chosen to achieve the intended risk levels. This choice is obviously not unique. Anderson (1960) approximates the OC and ASN-functions by transforming the problem of testing the mean of a normal distribution with known variance into a study of a Wiener process. The expression of the ASN-function, however, is too complicated to

allow for any useful analytic study of some optimal choice in C_5. Let us concentrate on the special case $\alpha = \beta$. Since the given hypotheses formulated in terms of a Wiener process approximation to the log likelihood ratio are symmetrically disposed around $(\lambda_1+\lambda_0)/2$ it would seem reasonable to concentrate on symmetrical procedures within C_5. For symmetrical procedures, sup $E(N;\lambda)$ occurs at $\lambda = (\lambda_1+\lambda_0)/2$. It is then possible numerically to find the symmetric procedure that satisfies $Q(\lambda_0) = 1-\alpha$ (which implies $Q(\lambda_1) = \alpha$) and minimizes $E(N; (\lambda_1+\lambda_0)/2)$. Gosh compares this test with two-sided truncated SPRT with respect to the ASN-function. Under the normal assumption of the log likelihood ratios based on Gosh's results, we could expect the two-sided SPRT to give almost the same value of the ASN-function at λ_0 and λ_1 and a somewhat higher maximum of the ASN-function (Gosh's results indicate that the increase should be less than 10 per cent, even for $\alpha = \beta = 0.01$). As the use of the two-sided SPRT facilitates the computation we have concentrated on this test.

Results arrived at by Gosh (1970, p. 231) can be used to find an expression for the one-sided SPRT. Introducing

$$\eta = \frac{(2\lambda-\lambda_0-\lambda_1)\sqrt{\lambda \ln(\lambda_1/\lambda_0)}}{2(\lambda_1-\lambda_0)}\ ;\ a' = \frac{a}{\sqrt{\lambda \ln(\lambda_1/\lambda_0)}}\ ;\ b' = \frac{b}{\sqrt{\lambda \ln(\lambda_1/\lambda_0)}} \qquad \text{(III.2.5)}$$

Result I (One-sided SPRT with an upper finite boundary; C_3)

For each restricted sequential test of $H_0; \lambda \le \lambda_0$ against $H_1: \lambda \ge \lambda_1 (\lambda_1 \ge \lambda_0)$ with the properties

i) $Q(\lambda_0) \ge 1-\alpha$ and $Q(\lambda_1) \le \beta$

ii) the maximal number of observations are N_0

iii) H_1 is accepted if $S_n \ge a$ (where $0 < a < \infty$) for some n such that $0 < n \le N_0$

iv) H_0 is accepted if $S_n < a$ for all $n \le N_0$

then for every value of λ $(0 < \lambda < \infty)$

$$P(N < n|\lambda) = \begin{cases} \phi(\eta\sqrt{n} - \frac{a'}{\sqrt{n}}) + e^{2a'\eta}\phi(-\eta\sqrt{n} - \frac{a'}{\sqrt{n}}) \text{ when } 0 < n < N_0 \\ 1 \text{ if } n \geq N_0 \end{cases} \quad \text{(III.2.6)}$$

$$Q(\lambda) = \phi(-\eta\sqrt{N_0} + \frac{a'}{\sqrt{N_0}}) - e^{2a'\eta}\phi(-\eta\sqrt{N_0} - \frac{a'}{\sqrt{N_0}}) \quad \text{(III.2.7)}$$

$$E(n|\lambda) \simeq N_0 + (\frac{a'}{\eta} - N_0)\phi(\eta\sqrt{N_0} - \frac{a'}{\sqrt{N_0}}) - (\frac{a'}{\eta} + N_0)e^{2a'\eta}\phi(-\eta\sqrt{N_0} - \frac{a'}{\sqrt{N_0}}). \quad \text{(III.2.8)}$$

Result II (One-sided SPRT with a lower finite boundary; C_4)

For each restricted sequential test of $H_0: \lambda \leq \lambda_0$ against $H_1: \lambda \geq \lambda_1$ $(\lambda_1 > \lambda_0)$ with the properties

i) $Q(\lambda_0) \geq 1-\alpha$ and $Q(\lambda_1) \leq \beta$

ii) the maximal number of observations are N_0

iii) H_0 is accepted if $S_n \leq b$ $(-\infty < b < 0)$ for some n such that $0 < n < N_0$

iv) H_1 is accepted if $S_n > b$ for all $n \leq N_0$

then for each λ $(0 < \lambda < \infty)$

$$P(N < n|\lambda) \simeq \begin{cases} \phi(\frac{b'}{\sqrt{n}} - \eta\sqrt{n}) + e^{2b'\eta}\phi(\frac{b'}{\sqrt{n}} + \eta\sqrt{n}) \text{ if } 0 < n < N_0 \\ 1 \text{ if } n \geq N_0 \end{cases} \quad \text{(III.3.9)}$$

$$Q(\lambda) \simeq \phi(\eta\sqrt{N_0} - \frac{b'}{\sqrt{N_0}}) - e^{2b'\eta}\phi(\eta\sqrt{N_0} + \frac{b'}{\sqrt{N_0}}) \quad \text{(III.3.10)}$$

$$E(N;\lambda) \simeq N_0 + (\frac{b'}{\eta} - N_0)\phi(-\eta\sqrt{N_0} + \frac{b'}{\sqrt{N_0}}) - (\frac{b'}{\eta} + N_0)e^{2b'\eta}\phi(\eta\sqrt{N_0} + \frac{b'}{\sqrt{N_0}}). \quad \text{(III.3.11)}$$

Expressions presented by Gosh (1970, p. 237) could also be used to find approximate expressions of the OC and ASN-functions of two-sided SPRT. However, computation of the boundaries a and b will be easier if we use another of Gosh's arguments (1970, p. 222). If the excess of S_N over a and b at termination of

the test is small, the solutions to equations (III.2.12) and (III.2.13) below guarantee minimum strength (α, β) for the truncated SPRT. If the parameters are defined as in (III.2.5), we obtain the following:

Result III (two-sided truncated SPRT)

For each truncated sequential test of $H_0: \lambda \le \lambda_0$ against $H_1: \lambda \ge \lambda$ with the properties

i) $Q(\lambda_0) \ge 1-\alpha$ and $Q(\lambda_1) \le \beta$

ii) the maximal number of observations are N_0

iii) H_0 or H_1 is accepted according to whether the lower of the upper boundary in the inequality $b < Z_n < a$ is violated first when $n < N_0$

iv) at N_0 accept H_0 if $S_{N_0} \ge 0$, otherwise accept H_1

then the constants a and b can be determined from the equations

$$\frac{e^{a-1}}{e^a - e^b} - P(0 \le Z_{N_0} < a \,|\, \lambda_0) \cong 1-\alpha \qquad \text{(III.2.12)}$$

$$\frac{e^b(e^a-1)}{e^a - e^b} + P(b < Z_{N_0} < 0 \,|\, \lambda_1) \cong \beta \qquad \text{(III.2.13)}$$

and the average sampling number would be given approximately by

$$E_{c,N_0}(N;\lambda) \cong N_0 - \sum_{k=1}^{4}\sum_{j=0}^{\infty} \left(\frac{\lambda_{kj}}{\eta} + N_0\right) \delta_k e^{\eta\lambda_{kj}} \phi\left(-\eta\sqrt{N_0} - \frac{\lambda_{kj}}{\sqrt{N_0}}\right) +$$

$$+ \sum_{k=1}^{4}\sum_{j=0}^{\infty} \left(\frac{\lambda_{kj}}{\eta} - N_0\right) \delta_k e^{-\eta\lambda_{kj}} \phi\left(\eta\sqrt{N_0} - \frac{\lambda_{kj}}{\sqrt{N_0}}\right). \qquad \text{(III.2.14)}$$

where

$$\delta_1 = -e^{b'\eta} = -\delta_2 \qquad \delta_3 = -e^{a'\eta} = -\delta_4$$

$$\lambda_{1j} = 2(j+1)a' - (2j+1)b' \qquad \lambda_{2j} = 2ja' - (2j+1)b'$$

$$\lambda_{3j} = (2j+1)a' - 2(j+1)b' \qquad \lambda_{4j} = (2j+1)a' - 2jb' .$$

The expression of the average sampling number of the two-sided truncated SPRT is derived in Appendix E, where also computational procedures for finding the

boundaries of the continuation region and the truncation point and for calculating the average sampling number are presented. For one-sided SPRT, there is only one equation to solve for N_0 as a is set equal to $\ln[(1-\beta)/\alpha]$ or $b = \ln[\beta/(1-\alpha)]$. For the two-sided test the optimal test in the sense of Anderson (1960), Kiefer and Weiss (1957) will be that which achieves the risks (α, β) and which minimizes $\sup E(N;\lambda)$. Choosing a symmetric procedure and assuming that the observations X_i are Poisson distributed, this maximum should occur at the point where E(Z) equals zero, that is $\lambda' = (\lambda_1-\lambda_0)/\ln(\lambda_1/\lambda_0)$. When the log likelihood ratios are normally distributed, as in the approximation here, this maximum occurs at $\lambda = (\lambda_1+\lambda_0)/2$. The problem then is to choose a and N_0 so that $E[N;(\lambda_1+\lambda_0)/2]$ is minimized. The limit of the expression of E(N), given in III.2.14, when $\eta \to 0$ tends to N_0, why no such test can be constructed here. Thus, the quality of tests with different truncation points will be studied instead.

III.3 Sequential Probability Ratio Tests for Comparing the Means of Two Poisson Distributed Random Variables

III.3.1 Matched Observations

Let (X_i, Y_i) be a sequence of mutually independent and identically distributed random variables. The single variables X_i and Y_i are assumed to be independent and distributed according to Poisson distributions with means λ_x and λ_y. The common joint p.m.f. of (X_i, Y_i) can now be written in the form

$$f(x,y;\gamma,\delta) = \frac{e^{-\gamma}\gamma^u}{u!}\left(\frac{u}{x}\right)\left(\frac{1}{1+\delta}\right)^x\left(\frac{\delta}{1+\delta}\right)^{u-x} \quad \text{for } x,y = 0,1,\ldots \qquad \text{(III.3.1.1)}$$

where

$$\gamma = \lambda_x + \lambda_y \qquad \delta = \frac{\lambda_y}{\lambda_x} \qquad u = x + y\,.$$

Gosh (1970) reasons as follows when he describes a SPRT for this situation. Since, for every n, the statistic $T_n = \Sigma(x_i+y_i)$ is Frazer sufficient for γ and

the family $[G_n(E_n|T_n, \delta),\ 0 < \delta < \infty]$ with

$$g_n(E_n|T_n, \delta) = [\delta/(1+\delta)]^{\sum_1^n y_i} \quad [1/(1+\delta)]^{\sum_1^n x_i}$$

has a monotone likelihood ratio in $V_n(E_n) = \sum_1^n x_i$, a conditional SPRT can be applied to test $H_0: \lambda_y/\lambda_x \le \delta_0$ against $H_1: \lambda_y/\lambda_x \ge \delta_1$. The SPRT S(a, b) observes $\{x_i, y_i\}$ successively and accepts or rejects H_0 according to whether the lower or the upper inequality in

$$b < \sum_1^n z_i < a \qquad \text{(III.3.1.2)}$$

is violated first. The test statistic Z_i is given by

$$Z_i = X_i \ln \frac{1+\delta_0}{1+\delta_1} + Y_i \ln \frac{\delta_1(1+\delta_0)}{\delta_0(1+\delta_1)} \quad \text{for } i \ge 1 \qquad \text{(III.3.1.3)}$$

and the constants a and b are chosen in such a way that

$$L(\delta_0) \cong 1-\alpha \quad \text{and} \quad L(\delta_1) \cong \beta \quad \text{for all } \gamma.$$

If S_n does not involve any excess over b and a, we get a UME SPRT with true risks equal to α and β independent of γ if we chose

$$b = \ln \frac{\beta}{1-\alpha} \quad \text{and} \quad a = \ln \frac{1-\beta}{\alpha}. \qquad \text{(III.3.1.4)}$$

If S_n has an excess over b and a, the true risks depend on γ. But as long as the excess of S_n over a and b and the risks (α, β) are small, the constants in (IV.1.4) give a nearly UME test and the true risks (α', β') satisfy

$$\alpha'(\gamma, \delta_0) \cong \alpha, \quad \beta'(\gamma, \delta,) \cong \beta \quad \text{and} \quad \alpha'(\gamma, \delta_0) + \beta'(\gamma, \delta,) < \alpha + \beta \quad \text{for all } \gamma.$$

As in the one-sample problem this test minimizes $E(N, \gamma, \delta_0)$ and $E(N, \gamma, \delta_1)$ uniformly in γ among all tests with risks (α, β) or less. We shall now here study truncated tests for this situation.

Let $\varphi(h; \gamma, \delta) = E(e^{hZ}; \gamma, \delta)$ be the moment-generating function of the random variable Z. We then have

$$E(e^{hZ}) = \exp\{-\gamma[1-\{\frac{1}{1+\delta}(\frac{1+\delta_0}{1+\delta_1})^h + \frac{\delta}{1+\delta}(\frac{\delta_1(1+\delta_0)}{\delta_0(1+\delta_1)})^h\}]\} . \qquad \text{(III.3.1.5)}$$

Writing $p = \frac{1}{1+\delta}$, $p_1 = \frac{1}{1+\delta_0}$, $p_0 = \frac{1}{1+\delta_1}$ the moment-generating function of Z can be written as

$$\varphi(h;\gamma,\delta) = \exp\{-\gamma[1-\{p(\frac{p_1}{p_0})^h + (1-p)(\frac{1-p_1}{1-p_0})^h\}]\} . \qquad \text{(III.3.1.6)}$$

It can be shown that $\varphi(h;\gamma,\delta)$ as a function of h is nonsingular at $h = 0$ and $h = h_0 \neq 0$, where h_0 is the non-zero root to the equation

$$\ln\varphi(h;\gamma,\delta) = 0 . \qquad \text{(III.3.1.7)}$$

For sufficiently small $|t|$ the equation

$$\ln\varphi(h;\gamma,\delta) = -it \qquad \text{(III.3.1.8)}$$

has two roots $h_1(it)$ and $h_2(it)$ such that

$$\lim_{t\Rightarrow 0} h_1(it) = 0 \qquad \lim_{t\Rightarrow 0} h_2(it) = h_0 ,$$

where h_0 is the non-zero solution to the equation $\ln\varphi(h;\gamma,\delta) = 0$. Here, this equation equals

$$\varphi_p(h) = p(\frac{p_1}{p_0})^h + (1-p)(\frac{1-p_1}{1-p_0})^h = 1 \qquad \text{(III.3.1.9)}$$

which is identical to the fundamental lemma applied to a Bernoulli variable. It has not been possible to express the solution to this equation analytically. However, as in the one-sample case, different analytical approximations of the exact roots can be studied.

Now write equation (III.3.1.8) as (cf. Section II.2)

$$K(h^2 - h_0 h) + it = 0 \qquad \text{(III.3.1.10)}$$

where K is chosen in an appropriate way.

By applying the fact that we should have $h_0(p_0) = 1$ and $h_0(p_1) = 1$ and assuming that $h_0(p)$ can be written as either $h_0(p) = a + bp$ or $h_0(p) = c + d/p$, two expressions h_{0I} and h_{0II} can be found which are equivalent to those for $h_0(\lambda)$ in Section II.2.

A third approximation h_{0T} can be found if $\exp\{h \ln(p_1/p_0)\}$ and $\exp\{h \ln[(1-p_1)/(1-p_0)]\}$ are expanded around $h = 0$ up to and including the terms of order h^2. A fourth expression, h_{0N}, will be found if the log likelihood ratio Z for testing the parameter p in a Bernoulli distribution is approximated by means of a normal distribution having the same mean and variance as Z. The four approximations will then be

$$h_{0I} = a + bp = \frac{p_1 + p_0 - 2p}{p_1 - p_0}$$

$$h_{0II} = c + \frac{d}{p} = \frac{2p_1p_0 - (p_1+p_0)p}{p(p_1 - p_0)}$$

(III.3.1.11)

$$h_{0T} = \frac{-2\left\{p \ln\left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right] + \ln\left(\frac{1-p_1}{1-p_0}\right)\right\}}{p \cdot \ln^2\left(\frac{p_1}{p_0}\right) + (1-p)\ln^2\left(\frac{1-p_1}{1-p_0}\right)}$$

$$h_{0N} = \frac{-2\left\{p \ln\left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right] + \ln\left(\frac{1-p_1}{1-p_0}\right)\right\}}{p(1-p)\ln^2\left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right]} .$$

Values for these approximations, the corresponding values of $\varphi_p(h)$ and the exact values of h and h_{min}, i.e. the value of h giving the minimum of $\varphi_p(h)$, have been calculated for different sets of p, p_1 and p_0 (cf. Section II.2). The exact values of h and h_{min} have been calculated using Simpson-Raphson's formula with a maximum difference between the last two iterations less than or equal to 0.01. For values of $p_1 = 0.50$, $p_0 = 0.40$, 0.45, 0.48 and

$p = p_0\{1+k(p_1-p_0)\}$ where $k = -5, -4, \ldots, 15$, Table III.3.1.1 shows that there is little difference between h_{0I} and h_{0T} but that a combination of h_{0I} and h_{0T} provides the best approximation, and usually a good approximation to the exact rott h_0. The approximation h_{0T} is better for the smallest values of p and for p around $(p_1+p_0)/2$, while h_{0I} is the better approximation for values of p near p_0 and p_1, and for large values of p. As h_{0I} gives the correct value of h_0 when $p = p_0$ or p_1 — which is important for deciding the truncation point in truncated tests — this approximation was chosen. In order to obtain more accurate values of the OC and AST-functions for truncated SPRT, we can try using a combination of h_{0I} and h_{0T}.

As the boundaries to the continuation region and the truncation point are determined from the distribution of N when p equals p_0 and p_1, we use h_{0I}. To find the characteristics of the decisive sampling number, N, for one-sided and two-sided truncated SPRT we can use the results in Section III.2.

Chose η as

$$\eta = \frac{h_0}{2}\sqrt{\frac{k}{2}} = \frac{(p_1+p_0-2p)}{2(p_1-p_0)}\sqrt{p\cdot\ln^2\left(\frac{p_1}{p_0}\right)+(1-p)\ln^2\left(\frac{1-p_1}{1-p_0}\right)}\,. \qquad \text{(III.3.1.12)}$$

Then introducing δ again write

$$k(\delta) = \sqrt{\frac{1}{1+\delta}\ln^2\left[\frac{1+\delta_0}{1+\delta_1}\right]+\frac{\delta}{1+\delta}\ln^2\left[\frac{\delta_1(1+\delta_0)}{\delta_0(1+\delta_1)}\right]} \qquad \text{(III.3.1.13)}$$

and

$$\eta = \frac{[2\delta-\delta_0(1+\delta_1)-\delta_1(1+\delta_0)+\delta(\delta_1+\delta_0)]\cdot k(\delta)}{2(1+\delta)(\delta_1-\delta_0)}\;;\; a' = \frac{a}{k(\delta)}\;;\; b' = \frac{b}{k(\delta)}\,.$$

The errors in the expressions in Section III.2 are due to two assumptions:

i) the approximation of the random walk for Z with a continuous time process with discrete states. In this case the error depends on the mean and variance of the test statistic Z and tends toward zero when $|E(Z)|$ and $V(Z)$ tends toward zero. Note that

$$E(z;\gamma,\delta) = \frac{\gamma}{(1+\delta)} \left\{\delta \ln \left[\frac{\delta_1(1+\delta_0)}{\delta_0(1+\delta_1)}\right] - \ln\left(\frac{1+\delta_1}{1+\delta_0}\right)\right\}$$

$$V(z;\gamma,\delta) = \frac{\gamma}{(1+\delta)} \left\{\delta \ln^2 \left[\frac{\delta_1(1+\delta_0)}{\delta_0(1+\delta_1)}\right] + \ln^2\left(\frac{1+\delta_1}{1+\delta_0}\right)\right\}. \qquad \text{(III.3.1.14)}$$

ii) the approximation of the continuous time process with discrete states to a Wicner process.

III.3.2 Comparison of a Fixed-size Sample with Sequential Observations

In some situations the effects of a treatment can be tested by studying the objects before and after the treatment has been applied (perhaps with a delay to eliminate disturbances). In this case the observations before treatment is a fixed-sample. Suppose that we have drawn m independent observations $(y_1, y_2, \ldots, y_m)$ from a random variable, Y, that follows a Poisson distribution with mean λ_y.

After the treatment has been applied, we take independent samples of the random variable, X, which is assumed to be independent of Y (a delay in the sampling makes this proposition reasonable) and have a Poisson distribution with mean λ_x. We now want to test the hypothesis $H_0: \lambda_x = \lambda_y = \lambda_0$ against the alternative $H_1: \lambda_x = \lambda_y(1+\Delta) = \lambda_1$ subject to the restrictions $Q(\lambda_0) \geq 1-\alpha$ and $Q(\lambda_1) \leq \beta$, where $Q(\lambda)$ is the operating characteristic function of the test. The probability distribution of an observation x and the sample $(y_1, \ldots, y_m)$ is

$$f(x, y_1, y_2 \ldots, y_m; \lambda_x, \lambda_y) \approx \frac{e^{-\lambda_x}(\lambda_x)^x e^{-m\lambda_y}(\lambda_y)^{\sum_1^m y_j}}{x! \prod_1^m y_j!}. \qquad \text{(III.3.2.1)}$$

A test which does not depend on λ_0 cannot be constructed for testing hypotheses H_0 against H_1. This can be seen from the fact that the test statistic

$$Z_i = \ln f(x_i, y_1, \dots y_m; \lambda_0, \lambda_0[1+\Delta]) - \ln f(x_i, y_1, y_2, \dots y_m; \lambda_0, \lambda_0) =$$

$$= x_i \ln(1+\Delta) - \Delta\lambda_0 \qquad \text{(III.3.2.2)}$$

depends on the unknown value of λ_0 .

We can now construct an approximate test of H_0 against H_1 by using the estimate

$$\bar{y} = \frac{\sum_m y_j}{m}$$

of λ_y, which give us the test statistic

$$Z_i^* = x_i \cdot \ln(1+\Delta) - \Delta\bar{y} \ . \qquad \text{(III.3.2.3)}$$

As the boundaries of the continuation region do not depend on λ_0 and as $\bar{y}$ tends in the limit toward λ_y when $m \to \infty$, this test will be asymptotically equivalent to the test based on Z_i. We now study the effects of errors in the estimate of λ_y on the true risks, the OC-curve and the ASN-curve. (Functions associated with the test statistic Z_i^* are denoted by an asterisk.) Define the stochastic variable ϵ from the relations $\bar{y} = \lambda_0(1+\epsilon)$. The test statistic Z_i^* can then be written as

$$Z_i^* = x_i \cdot \ln(1+\Delta) - \Delta\lambda_y - \Delta\epsilon\lambda_y \qquad \text{(III.3.2.4)}$$

Let the moment-generating function of Z_i^* conditioned by ϵ be

$$\varphi_\lambda^*(h|\epsilon) = \varphi_\lambda(h) \cdot e^{-\Delta\epsilon\lambda} \ . \qquad \text{(III.3.2.5)}$$

The solution, h_0, to the equation

$$\ln \varphi_\lambda^*(h|\epsilon) = \ln \varphi_\lambda(h) - \Delta\epsilon\lambda_0 = 0 \qquad \text{(III.3.2.6)}$$

is

an increasing function of ϵ when $E(z) < 0$, that is $h(\lambda) > 0$

a decreasing function of ϵ when $E(z) > 0$, that is $h(\lambda) < 0$.

For the operating characteristic function, $Q(\lambda)$, of the test statistic Z we have

$$Q(\lambda) = \frac{e^{ah(\lambda)} - 1}{e^{ah(\lambda)} - e^{bh(\lambda)}} \qquad \text{(III.3.2.7)}$$

where $h(\lambda)$ is the non-zero solution to $\ln \varphi_{\lambda}(h) = 0$. $Q(\lambda)$ is an increasing function of $h(\lambda)$. Then, the conditional operating characteristic function of the test statistic Z^* conditioned by ϵ, $Q(\lambda|\epsilon)$ is

an increasing function of ϵ, when $E(z) < 0$

a decreasing function of ϵ, when $E(z) > 0$.

For the average sampling number using the test statistic Z, we have approximately

$$E(N;\lambda) = \frac{b \cdot Q(\lambda) + a[1 - Q(\lambda)]}{E(z)} \qquad \text{(III.3.2.8)}$$

This function is

a decreasing function of $h(\lambda)$ when $E(z) < 0$

an increasing function of $h(\lambda)$ when $E(z) < 0$.

Then, the conditional average sampling number using the test statistic Z^* conditioned by ϵ, $E(N;\lambda|\epsilon)$ is a decreasing function of ϵ.

If we now chose ϵ in such a way that $P\{(1-\epsilon)\lambda_y < \bar{y} < (1+\epsilon)\lambda_y\} = p$ the following relations are obtained:

$$P\{Q^*[\lambda_x|\bar{y} = (1-\epsilon)\lambda_0] < Q^*(\lambda_x|\bar{y}) < Q^*[\lambda_x|\bar{y} = (1+\epsilon)\lambda_0]\} \geq p \qquad \text{(III.3.2.9)}$$

$$P\{E^*[N;\lambda_x|\bar{y} = (1+\epsilon)\lambda_0] < E^*(N;\lambda_x|\bar{y}) < E^*[N;\lambda_x|\bar{y} = (1-\epsilon)\lambda_0]\} \geq p \; . \qquad \text{(III.3.2.10)}$$

If we define $R(\lambda_x|\bar{y})$ as

$$R(\lambda_x|\bar{y}) = \frac{E^*(N;\lambda_x|\bar{y})}{E(N;\lambda_x)} \; , \qquad \text{(III.3.2.11)}$$

then the relation (III.3.2.10) can be written as

$$P\left\{\left|\frac{E^*(N;\lambda_x|\bar{y})}{E(N;\lambda_x)} - 1\right| > \max\ [1 - R(\lambda_x|\bar{y} = (1-\varepsilon)\lambda_0)]\ ,\right.$$

$$\left.[1 - R(\lambda_x|\bar{y} = (1+\varepsilon)\lambda_0)]\right\} \leq p\ . \qquad \text{(III.3.2.12)}$$

By using the fact that $V(\bar{y}) = \lambda_0/m$, we can now find a minimum value of m necessary to assure that $Q^*(\lambda_x|\bar{y})$ and $E^*(N;\lambda_x|\bar{y})$ differs from the true values $Q(\lambda_x)$ and $E(N;\lambda_x)$ mostly in a given proportion with a probability less than or equal to p.

Intervals for the maximal error in the risklevels for different values of ε, Δ and λ_0, when the intended risk levels are $\alpha = \beta = 0.05$, are shown in Table III.3.2. This table indicates that the errors increase with ε, Δ and λ_0. Table III.3.3 shows values of the ratio R for the different values of ε, Δ and λ_0 that will be achieved when the relative error in the estimate equals ε. As can be seen from the table, the maximal relative error in the average sampling numbers $E(N;\lambda_0)$ and $E(N;\lambda_1)$ seems to be only slightly dependent on λ_0, increases with ε and decreases with Δ. The results in these two tables are combined in Table III.3.1 which contains, for values of ε, the regions for Δ and λ_0 when the relative errors in risks and average sampling numbers $E(N;\lambda_0)$ and $E(N;\lambda_1)$ are controlled. The first column and the last three columns indicate the values of Δ and λ_0 for which the maximal error in the risk levels are smaller than 5, 10 and 25 per cent, respectively, of the intended risks 0.05. The third column shows tha maximal relative error in the average sampling number at λ_0 and λ_1 which can be ascertained at the same time — increased to the nearest 5 per cent level. It can then be seen from Table III.3.1 that e.g. when $\Delta = 0.50$, $\varepsilon = 0.02$, a maximal 10 per cent relative change in the risk levels can only be assured for values of λ_0 smaller than or equal to 1.0.

If we turn to Table III.3.2c we see that when λ_0 equals 1.0 the maximal relative error is about 4 per cent and when λ_0 equals 5.0 it is about 25 per cent, while for $\lambda_0 = 10.0$ the maximal relative error is about 50 per cent and for $\lambda_0 = 15.0$

it is about 85 per cent. It then seems reasonable that if λ_0 increases up to about 2.0 the relative error in the risk levels would not be much larger than 10 per cent. The third column indicates that the relative error in the average sampling numbers when λ equals λ_0 or λ_1 will be smaller than 10 per cent.

Table III.3.1 Comparison of a fixed-size sample with sequential observations
Values of ϵ and λ_0 for which errors in the risk levels and the average sampling number for $\lambda = \lambda_0$ and λ_1 are of given size, when the intended risk levels are $\alpha = \beta = 0.05$ and ϵ is the relative error in the estimate $\bar{y}$ of the value λ_0.

Δ given by $\lambda_1 = (1+\Delta)\lambda_0$	Relative error in estimate of λ_0 ϵ	Relative maximal error in $E(N;\lambda_0)$ and $E(N;\lambda_1)$ are smaller than	Values of λ_0 for which Maximal relative error in intended risk levels are less than or equal to		
			5 %	10 %	25 %
0.10	0.01	0.25	$\lambda_0 \leq 15.0$	$\lambda_0 \leq 15.0$	$\lambda_0 \leq 15.0$
0.25	0.01	0.10	$\lambda_0 \leq 5.0$	$\lambda_0 \leq 10.0$	$\lambda_0 \leq 10.0$
	0.02	0.25	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 5.0$	$\lambda_0 \leq 15.0$
0.50	0.01	0.05	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 5.0$	$\lambda_0 \leq 15.0$
	0.02	0.10	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 5.0$
	0.03	0.20	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 5.0$
	0.04	0.25	-	$\lambda_0 \leq 1.0$	$\lambda_0 \leq 1.0$

Table III.3.2.a Comparison of a fixed-size sample with sequential observations
Intervals for the change in the actual risk levels for different values of λ_0 and ε. (Cf Section III.3.2.) Intended risk levels $\alpha = \beta = 0.05$. $\Delta = 0.10$

Relative error in $\frac{\bar{y}}{\varepsilon}$	Intervals for actual risk levels for λ_0, when λ_0							
	1.00		5.00		10.00		15.00	
0.01	-0.0001	0.0001	-0.0007	0.0007	-0.0014	0.0014	-0.0021	0.0021
0.02	-0.0003	0.0003	-0.0014	0.0014	-0.0029	0.0029	-0.0044	0.0044
0.03	-0.0004	0.0004	-0.0021	0.0021	-0.0044	0.0044	-0.0067	0.0067
0.04	-0.0006	0.0006	-0.0029	0.0029	-0.0059	0.0059	-0.0091	0.0091
0.05	-0.0007	0.0007	-0.0036	0.0036	-0.0075	0.0075	-0.0116	0.0116
0.10	-0.0014	0.0014	-0.0065	0.0075	-0.0123	0.0160	-0.0173	0.0257
0.15	-0.0021	0.0021	-0.0095	0.0116	-0.0173	0.0257	-0.0236	0.0426
0.20	-0.0027	0.0029	-0.0123	0.0160	-0.0216	0.0366	-0.0287	0.0629
0.25	-0.0034	0.0036	-0.0149	0.0207	-0.0254	0.0490	-0.0329	0.0870
0.30	-0.0040	0.0044	-0.0173	0.0257	-0.0287	0.0629	-0.0362	0.1153

Table III.3.2.b Comparison of a fixed-size sample with sequential observations
Intervals for the change in the actual risk levels for different values of λ_0 and ε. (Cf Section III.3.2.) Intended risk levels $\alpha = \beta = 0.05$. $\Delta = 0.25$

Relative error in $\frac{\bar{y}}{\varepsilon}$	Intervals for actual risk levels for λ_0, when λ_0							
	1.00		5.00		10.00		15.00	
0.01	-0.0003	0.0004	-0.0017	0.0018	-0.0034	0.0036	-0.0050	0.0055
0.02	-0.0007	0.0007	-0.0034	0.0036	-0.0065	0.0075	-0.0095	0.0115
0.03	-0.0010	0.0011	-0.0050	0.0055	-0.0095	0.0116	-0.0136	0.0183
0.04	-0.0014	0.0014	-0.0065	0.0075	-0.0123	0.0160	-0.0173	0.0257
0.05	-0.0017	0.0018	-0.0081	0.0095	-0.0149	0.0207	-0.0206	0.0338
0.10	-0.0034	0.0036	-0.0149	0.0207	-0.0254	0.0490	-0.0329	0.0870
0.15	-0.0050	0.0055	-0.0206	0.0338	-0.0329	0.0870	-0.0401	0.1662
0.20	-0.0065	0.0075	-0.0254	0.0490	-0.0381	0.1366	-0.0494	0.2739
0.25	-0.0081	0.0095	-0.0295	0.0667	-0.0417	0.1990	-0.0467	0.4041
0.30	-0.0095	0.0160	-0.0329	0.0870	-0.0442	0.2739	-0.0481	0.5410

Table III.3.2.c Comparison of a fixed-size sample with sequential observations
Intervals for the change in the actual risk levels for different values of λ_0 and ϵ. (Cf Section III.3.2.) Intended risk levels $\alpha = \beta = 0.05$. $\Delta = 0.50$

Relative error in $\frac{\bar{y}}{\epsilon}$	Intervals for actual risk levels for λ_0, when λ_0 = 1.00		5.00		10.00		15.00	
0.01	-0.0007	0.0007	-0.0034	0.0036	-0.0065	0.0075	-0.0095	0.0116
0.02	-0.0014	0.0014	-0.0065	0.0075	-0.0123	0.0160	-0.0173	0.0257
0.03	-0.0021	0.0021	-0.0095	0.0116	-0.0173	0.0257	-0.0236	0.0426
0.04	-0.0027	0.0029	-0.0123	0.0160	-0.0216	0.0366	-0.0287	0.0629
0.05	-0.0034	0.0036	-0.0149	0.0207	-0.0254	0.0490	-0.0329	0.0870
0.10	-0.0065	0.0075	-0.0254	0.0490	-0.0381	0.1366	-0.0442	0.2739
0.15	-0.0095	0.0116	-0.0329	0.0870	-0.0442	0.2739	-0.0481	0.5410
0.20	-0.0123	0.0160	-0.0381	0.1366	-0.0472	0.5410	-0.0494	0.7634
0.25	-0.0149	0.0207	-0.0417	0.1990	-0.0487	0.6261	-0.0498	0.8793
0.30	-0.0173	0.0257	-0.0442	0.2739	-0.0494	0.7634	-0.0499	0.9254

Table III.3.3.a Comparison of a fixed-size sample with sequential observations
Intervals for the ratio between $E^*(N;\lambda|\bar{y})$ and $E(N;\lambda)$ when λ equals λ_0 and λ_1 for different values of λ_0 and ϵ. Intended risk levels $\alpha = \beta = 0.05$. $\Delta = 0.10$

Relative error in $\frac{\bar{y}}{\epsilon}$	Maximal intervals for the ratio when λ_0 = 1.00		5.00		10.00		15.00	
0.01	0.82	1.27	0.83	1.27	0.83	1.27	0.83	1.26
0.02	0.70	1.74	0.70	1.74	0.71	1.74	0.71	1.73
0.03	0.61	2.77	0.61	2.76	0.61	2.75	0.62	2.73
0.04	0.54	6.79	0.54	6.75	0.55	6.71	0.55	6.66
0.05	0.48		0.49		0.49		0.49	
0.10	0.32		0.32		0.33		0.33	
0.15	0.24		0.24		0.25		0.25	
0.20	0.19		0.20		0.20		0.20	
0.25	0.16		0.16		0.16		0.17	
0.30	0.14		0.14		0.14		0.15	

Where no values are given in Table III.3.3 it means that the calculated values are extremely high or inaccurate.

Table III.3.3.b Comparison of a fixed-size sample with sequential observations
Intervals for the ratio between $E^*(N;\lambda|\bar{y})$ and $E(N;\lambda)$ when λ equals λ_0 and λ_1 for different values of λ_0 and ε. Intended risk levels $\alpha = \beta = 0.05$.
$\Delta = 0.25$

Relative error in $\bar{y}$ smaller than ε	Maximal intervals for the ratio when λ_0							
ε	1.00		5.00		10.00		15.00	
0.01	0.92	1.10	0.92	1.10	0.92	1.09	0.92	1.09
0.02	0.84	1.23	0.85	1.22	0.86	1.21	0.86	1.20
0.03	0.78	1.38	0.79	1.37	0.80	1.35	0.81	1.33
0.04	0.73	1.59	0.74	1.57	0.75	1.54	0.76	1.50
0.05	0.69	1.86	0.69	1.83	0.70	1.78	0.71	1.73
0.10	0.52	14.35	0.53	13.80	0.55	12.89	0.56	11.67
0.15	0.42		0.44		0.45		0.45	
0.20	0.35		0.37		0.38		0.38	
0.25	0.31		0.32		0.33		0.33	
0.30	0.27		0.28		0.29		0.29	

Table III.3.3.c Comparison of a fixed-size sample with sequential observations
Intervals for the ratio between $E^*(N;\lambda|\bar{y})$ and $E(N;\lambda)$ when λ equals λ_0 and λ_1 for different values of λ_0 and ε. Intended risk levels $\alpha = \beta = 0.05$.
$\Delta = 0.50$

Relative error in $\bar{y}$ smaller than ε	Maximal intervals for the ratio when λ_0							
ε	1.00		5.00		10.00		15.00	
0.01	0.95	1.05	0.96	1.05	0.96	1.04	0.97	1.03
0.02	0.91	1.11	0.92	1.10	0.93	1.08	0.94	1.05
0.03	0.87	1.18	0.88	1.16	0.90	1.12	0.91	1.08
0.04	0.83	1.26	0.85	1.22	0.87	1.17	0.88	1.09
0.05	0.80	1.35	0.82	1.30	0.84	1.21	0.85	1.10
0.10	0.66	2.09	0.69	1.89	0.71	1.48	0.72	
0.15	0.57	4.71	0.60	3.90	0.61	1.89	0.62	
0.20	0.50		0.53		0.54		0.54	
0.25	0.44		0.47		0.48		0.48	
0.30	0.40		0.42		0.43		0.43	

Where no values are given in Table III.3.3 it means either that the calculated values are extremely high or inaccurate.

Table III.3.1.1 Study of functions as approximations to the nonzero root of the equation $\varphi_p(h) = 1$

$$h_{0I} = a + bp = \frac{p_1 + p_0 - 2p}{p_1 - p_0}$$

$$h_{0II} = c + \frac{d}{p} = \frac{2p_1p_0 + (p_1 + p_0)p}{p(p_1 - p_0)}$$

$$h_{0T} = \frac{-2\left\{p \ln \left[\frac{p_1(1-p_0)}{p_0(1-p_1)}\right] + \ln\left(\frac{1-p_1}{1-p_0}\right)\right\}}{p \cdot \ln^2\left(\frac{p_1}{p_0}\right) + (1-p)\ln^2\left(\frac{1-p_1}{1-p_0}\right)}$$

$$h_{0N} = \frac{-2\left\{p \ln \frac{p_1(1-p_0)}{p_0(1-p_1)} + \ln\left(\frac{1-p_1}{1-p_0}\right)\right\}}{p(1-p)\ln^2 \frac{p_2(1-p_0)}{p_0(1-p_1)}}$$

$p_0 = 0.4$ $p_1 = 0.5$

p	h_{0I}	$\varphi_p(h_{0I})$	h_{0II}	$\varphi_p(h_{0II})$	h_T	$\varphi_p(h_T)$	h_N	$\varphi_p(h_N)$	h_{min}	h_0
0.350	2.0000	0.9983	2.4286	1.0192	2.0704	1.0012	10.6868	3.8921	1.0284	2.0427
0.360	1.8000	0.9989	2.1111	1.0122	1.8548	1.0009	9.4985	3.1078	0.9207	1.8301
0.370	1.6000	0.9994	1.8108	1.0071	1.6410	1.0007	8.3370	2.5154	0.8143	1.6197
0.380	1.4000	0.9997	1.5263	1.0036	1.4290	1.0005	7.2131	2.0666	0.7091	1.4114
0.390	1.2000	0.9999	1.2564	1.0013	1.2188	1.0003	6.1179	1.7273	0.6049	1.2049
0.400	1.0000	1.0000	1.0000	1.0000	1.0103	1.0002	5.0479	1.4728	0.5017	1.0000
0.410	0.8000	1.0001	0.7561	0.9994	0.8035	1.0001	3.9997	1.2855	0.3994	0.7966
0.420	0.6000	1.0001	0.5238	0.9992	0.5984	1.0000	2.9704	1.1524	0.2978	0.5944
0.430	0.4000	1.0001	0.3023	0.9994	0.3950	1.0000	1.9569	1.0644	0.1968	0.3932
0.440	0.2000	1.0000	0.0909	0.9998	0.1933	1.0000	0.9565	1.0151	0.0965	0.1931
0.450	0.0000	1.0000	-0.1111	1.0002	-0.0068	1.0000	-0.0334	1.0000	-0.0034	-0.0068
0.460	-0.2000	1.0000	-0.3043	1.0006	-0.2052	1.0000	-1.0154	1.0166	-0.1028	-0.2059
0.470	-0.4000	1.0000	-0.4894	1.0008	-0.4021	1.0000	-1.9919	1.0634	-0.2020	-0.4046
0.480	-0.6000	1.0000	-0.6667	1.0009	-0.5974	0.9999	-2.9653	1.1406	-0.3009	-0.6031
0.490	-0.8000	1.0000	-0.8367	1.0006	-0.7911	0.9998	-3.9379	1.2491	-0.3996	-0.8015
0.500	-1.0000	1.0000	-1.0000	1.0000	-0.9832	0.9997	-4.9122	1.3915	-0.4983	-1.0000
0.510	-1.2000	1.0000	-1.1569	0.9990	-1.1739	0.9994	-5.8903	1.5712	-0.5970	-1.1988
0.520	-1.4000	1.0001	-1.3077	0.9976	-1.3630	0.9990	-6.8748	1.7933	-0.6957	-1.3981
0.530	-1.6000	1.0001	-1.4528	0.9958	-1.5507	0.9985	-7.8679	2.0644	-0.7946	-1.5979
0.540	-1.8000	1.0001	-1.5926	0.9935	-1.7368	0.9979	-8.8722	2.3934	-0.8938	-1.7985

$p_0 = 0.45$ $p_1 = 0.50$

p	h_{0I}	$\varphi_p(h_{0I})$	h_{0II}	$\varphi_p(h_{0II})$	h_T	$\varphi_p(h_T)$	h_N	$\varphi_p(h_N)$	h_{min}	h_0
0.425	2.0000	0.9999	2.1765	1.0018	2.0170	1.0001	9.0323	1.3438	1.0068	2.0103
0.430	1.8000	0.9999	1.9302	1.0012	1.8133	1.0001	8.1044	1.2732	0.9050	1.8073
0.435	1.6000	1.0000	1.6897	1.0007	1.6100	1.0000	7.1833	1.2121	0.8035	1.6048
0.440	1.4000	1.0000	1.4545	1.0004	1.4071	1.0000	6.2686	1.1598	0.7022	1.4028
0.445	1.2000	1.0000	1.2247	1.0001	1.2046	1.0000	5.3595	1.1157	0.6012	1.2013
0.450	1.0000	1.0000	1.0000	1.0000	1.0026	1.0000	4.4556	1.0793	0.5004	1.0001
0.455	0.8000	1.0000	0.7802	0.9999	0.8009	1.0000	3.5562	1.0501	0.3999	0.7993
0.460	0.6000	1.0000	0.5652	0.9999	0.5007	1.0000	2.6608	1.0279	0.2995	0.5988
0.465	0.4000	1.0000	0.3548	0.9999	0.3989	1.0000	1.7689	1.0123	0.1992	0.3984
0.470	0.2000	1.0000	0.1489	1.0000	0.1984	1.0000	0.8797	1.0030	0.0992	0.1986
0.475	0.0000	1.0000	-0.0526	1.0000	-0.0016	1.0000	-0.0071	1.0000	-0.0008	-0.0093
0.480	-0.2000	1.0000	-0.2500	1.0001	-0.2012	1.0000	-0.8921	1.0031	-0.1007	-0.2013
0.485	-0.4000	1.0000	-0.4433	1.0001	-0.4005	1.0000	-1.7759	1.0122	-0.2005	-0.4010
0.490	-0.6000	1.0000	-0.6327	1.0001	-0.5993	1.0000	-2.6589	1.0274	-0.3002	-0.6007
0.495	-0.8000	1.0000	-0.8182	1.0001	-0.7977	1.0000	-3.5419	1.0486	-0.3999	-0.8003
0.500	-1.0000	1.0000	-1.0000	1.0000	-0.9957	1.0000	-4.4251	1.0760	-0.4996	-1.0000
0.505	-1.2000	1.0000	-1.1782	0.9999	-1.1934	1.0000	-5.3093	1.1097	-0.5992	-1.1997
0.510	-1.4000	1.0000	-1.3529	0.9997	-1.3906	0.9999	-6.1948	1.1499	-0.6989	-1.3996
0.515	-1.6000	1.0000	-1.5243	0.9994	-1.5875	0.9999	-7.0824	1.1968	-0.7987	-1.5995
0.520	-1.8000	1.0000	-1.6923	0.9991	-1.7840	0.9999	-7.9723	1.2507	-0.8984	-1.7996

III.4 Sequential Probability Ratio Tests of the Parameters in a Negative Binomial Distribution

One way to extend the Poisson model is to assume that the parameter λ is a random variable with distribution function $G(\lambda)$. If $G(\lambda)$ is a gamma distribution as is often used, that is having the density function

$$g(\lambda) = \frac{\mu^{\delta}}{\Gamma(\delta)} \lambda^{\delta-1} e^{-\mu\lambda} , \qquad \text{(III.4.1)}$$

the distributions of X reduces to the negative binomial distribution with

$$f(x; \mu, \delta) = \frac{\Gamma(x+\delta)}{x!\,\Gamma(\delta)} \quad \frac{\mu^{\delta}}{(1+\mu)^{x+\delta}} \quad \text{for } x = 0,1,2 \ldots \qquad \text{(III.4.2)}$$

The main purpose of deriving tests of the parameter μ in this distribution is to study the effects of the heterogeneity of λ on the OC and ASN-functions. This will be dealt with later on (cf Chapter VI)). However, practical use of the negative binomial distribution appears to be rather limited.

The number of events, X, during a given time interval $[0,T]$ will also have a negative binomial distribution if they follow a non-stationary birth process where the transition probabilities depend on time, i.e. a Polya process. If events have occurred at time t , then the probability of another event in the time interval $(t, t+h)$ equals

$$\lambda_h(t) \cdot h + 0(h)$$

where

$$\lambda_h(t) = \frac{x+\zeta}{t+\zeta} .$$

Neyman (1957 p. 7) discusses these two models and states "that the difference between the two models becomes apparent when the observations include something more than the variable X — the number of events during $[0, T]$ — perhaps the numbers X_1 and $X_2 = X - X_1$ of events incurred in two successive subintervals, the union of which constitutes the original period of observations". This problem of identifiability has to be taken into consideration when applying sequential tests.

The SPRT(a,b) of $H_0: \mu \leq \mu_0$ against $H_1: \mu \geq \mu_1$ when δ is known observes $\{x_i\}$ successively and accepts or rejects H_0 according to whether the lower or the upper inequality in

$$b < \sum_{1}^{n} z_i < a \qquad \text{(III.4.3)}$$

is violated first. The test statistic is

$$Z_i = \delta \ln \{\frac{\mu_1 (1+\mu_0)}{\mu_0 (1+\mu_1)}\} - X \ln (\frac{1+\mu_1}{1+\mu_0}) \qquad \text{(III.4.4)}$$

and the constants a and b are chosen in such a way that

$Q(\mu_0) \simeq 1-\alpha$ and $Q(\mu_1) \simeq \beta$.

In order to find the moment-operating function of the test statistic Z, we write

$$Z = \delta \cdot \ln \{\frac{\mu_1 (1+\mu_0)}{\mu_0 (1+\mu_1)}\} + X \ln (\frac{1+\mu_0}{1+\mu_1}) = \delta \cdot \ln f + x \ln g \; . \qquad \text{(III.4.5)}$$

The moment-generating function $\phi_\mu (h;\delta)$ is then given by

$$\varphi_\mu (h;\delta) = E(e^{hZ}) = e^{h\delta \cdot \ln f} E_\lambda E(e^{hx \cdot \ln g} | \lambda) =$$

$$= e^{h\delta \cdot \ln f} E_\lambda \{ e^{\lambda (1 - e^{h \cdot \ln g})} \} = \{\frac{\mu \cdot e^{h \cdot \ln f}}{1+\mu-e^{h \cdot \ln g}}\}^{\delta} \qquad \text{(III.4.6)}$$

If p is now identified as $\mu/(1+\mu)$, we find

$$\varphi_\mu (h;\delta) = \left[\frac{p \cdot e^{h \ln f}}{1-(1-p) \cdot e^{h \ln g}}\right]^{\delta} = \left[\frac{p (\frac{p_1}{p_0})^h}{1-(1-p)(\frac{1-p_1}{1-p_0})^h}\right]^{\delta} . \qquad \text{(III.4.7)}$$

To find the roots $h_1(it)$ and $h_2(it)$ to the equation

$$\ln \varphi_\mu (h;\delta) = -it \qquad \text{(III.4.8)}$$

the moment generating function of Z is approximated using a second-degree polynomial in h (cf Section III.3) and we get the equation

$$\tfrac{1}{2}\{p \cdot \ln^2 (\frac{p_1}{p_0}) + (1-p) \ln^2 (\frac{1-p_1}{1-p_0})\} [h^2 - h_0 h] + it = 0 \qquad \text{(III.4.9)}$$

where h_0 is the solution to the equation $\varphi_\mu (h;\delta) = 1$, which is identical to

$$p(\frac{p_1}{p_0})^h + (1-p)(\frac{1-p_1}{1-p_0})^h = 1 \ . \qquad \text{(III.4.10)}$$

This implies that the problem is transformed into tests of the parameter p in a Bernoulli distribution. We studied the same problem in Section III.3.1 and found that

$$h_0 = \frac{p_1 + p_0 - 2p}{p_1 - p_0} = \frac{(1+\mu_0)(\mu_1 - \mu) + (1+\mu_1)(\mu_0 - \mu)}{(1+\mu)(\mu_1 - \mu_0)} \ . \qquad \text{(III.4.11)}$$

writing

$$k(\mu) = \sqrt{\frac{\mu}{(1+\mu)} \ln^2 \left[\frac{\mu_1 (1+\mu_0)}{\mu_0 (1+\mu_1)}\right] + \frac{1}{1+\mu} \ln^2 \left[\frac{1+\mu_0}{1+\mu_1}\right]} \qquad \text{(III.4.12)}$$

and

$$\eta = \frac{\{(1+\mu_0)(\mu_1 - \mu) + (1+\mu_1)(\mu_0 - \mu)\} \cdot k(\mu)}{2(1+\mu)(\mu_1 - \mu_0)} \ ; \ a' = \frac{a}{k(\mu)} \ ; \ b' = \frac{b}{k(\mu)} \ ,$$

the results in Section III.2 can be used to give the characteristics of the decisive sampling number N for one-sided and two-sided SPRT. The error in the approximations will depend on δ, through E(z) and V(z). Furthermore, the error in the approximations will tend toward zero when $|E(z)|$ and V(z) tend toward zero.

III.5 Sequential Probability Ratio Tests of the Parameter in a Multiplicative Poisson Model

Another way to extend the Poisson model when studying the effects of a treatment is to use the multiplicative Poisson model. In this model pairwise observations $\{x_i, y_i\}$ are studied which are usually observations of the same sampling unit before and after the treatment has been applied. If we want to investigate whether the treatment has had any effect, a common procedure can be used whereby it is assumed that the conditional distribution of X, given the

value of $X + Y$, is binomial with a parameter p, that is

$$P(X = x \mid X + Y = m) = \binom{m}{x} p^x (1-p)^{m-x} \quad . \tag{III.5.1}$$

It can be shown that when this model is assumed, X_i and Y_i have Poisson distributions with parameters $p\lambda_i$ and $(1-p)\lambda_i$ where λ_i is arbitrary (cf. Moran (1952) and Chatterji (1963)). If the sets $\{X_i\}$ and $\{Y_i\}$ of random variables are assumed to be independent and have Poisson distributions with parameters $p\lambda_i$ and $(1-p)\lambda_i$, i.e. the effect of the treatment is proportionally constant for all sampling units, we find that the conditional distribution of X_i, given $X_i + X_i = m_i$, is the binomial distribution. As the parameters λ_i disappear in the conditional distribution, tests regarding hypotheses about the quotient between two sets of Poisson distributed variables can be constructed without knowledge of the λ_i's. Since the conditional distribution is the same only in the case where all sums $X_i + Y_i$ are the same, no SPRT can directly be constructed. Instead, the fact that if we rely on the probability that a given event which has occurred (belonging to the period before treatment) is p, a SPRT can be constructed using the Bernoulli distribution.

We get x_i ones and y_i zeros for each value of i and we can examine the observations in some natural order. Thus, we have m_i observations on the Bernoulli distribution with the p.m.f.

$$P(U = u \mid p) = p^u (1-p)^{1-u} \qquad u = 0, 1 \, .$$

The SPRT of $H_0: p \le p_0$ against $H_1: p \ge p_1$ observes $\{x_i; y_i\}$ successively and accepts or rejects H_0 according to whether the lower or the upper inequality in

$$b < \sum_{1}^{n} z_\ell < a \tag{III.5.2}$$

is violated first. The test statistic Z_ℓ is

$$Z_\ell = U \cdot \ln \left[\frac{p_1(1-p_0)}{p_0(1-p_1)} \right] + \ln \left[\frac{1-p_1}{1-p_0} \right] \tag{III.5.3}$$

where the constants a and b are chosen in such a way that $Q(p_0) = 1-\alpha$ and $Q(p_1) = \beta$.

The decisive sampling number N will now denote the accumulated number of events in the sequence $\{x_i, y_i\}$ and the number M of observations necessary to terminate the test will satisfy the inequality

$$\sum_{1}^{M-1} m_i < N \leq \sum_{1}^{M} m_i .$$

For the average sampling number E(M) of observations on $\{x_i, y_i\}$ equivalent to E(N) we have that $E(M) = E(N)E(\frac{1}{\lambda})$.

If we use $E(N)/\bar{\lambda}$, where $\bar{\lambda}$ is an estimate of the expected intensity based on observations, as an estimate of E(M), E(M) will be underestimated if all λ_i are not equal. In order to find a consistent estimate of E(M) we have to study the time t_i between events i and i + 1 and then estimate $E(\frac{1}{\lambda})$ by means of $(\sum_{1}^{n} t_i)/n$.

The moment-generating function $\varphi_p(h) = E(e^{hZ})$ of the variable Z is

$$\varphi_p(h) = p\left(\frac{p_1}{p_0}\right)^h + (1-p)\left(\frac{1-p_1}{1-p_0}\right)^h . \qquad \text{(III.5.4)}$$

A statistic that was directly related to the Bernoulli distribution was also studied in Section III.3.1. By writing

$$k(p) = \sqrt{p \cdot \ln^2\left(\frac{p_1}{p_0}\right) + (1-p) \cdot \ln^2\left(\frac{1-p_1}{1-p_0}\right)} \; ; \; \eta = \frac{p_1 + p_0 - 2p}{2(p_1 - p_0)} k(p) \; ;$$

$$a' = \frac{a}{k(p)} \; ; \; b' = \frac{b}{k(p)} \qquad \text{(III.5.5)}$$

the results for one-sied and two-sided SPRT presented in Section III.2 can be used.

Chapter IV EVALUATION OF APPROXIMATE EXPRESSIONS OF TEST CHARACTERISTICS

IV.1 Introduction

A study of the decisive sampling number, N, for the sequential probability ratio test of hypotheses $H_0: \lambda \leq \lambda_0$ against $H_1: \lambda \geq \lambda_1$ involves studying the number of observations up to and including the time when a discrete time process with discrete state space — where for each time, the observations are Poisson distributed with parameter λ — will pass either of two bounds a and b.

This approximation is made in two steps. In the first step the discrete time process is approximated to a continuous time process by neglecting the excess of the cumulative log likelihood ratio, S_N, over a and b at the termination of the test (cf. Wald (1947)). In the next step the discrete state space is approximated to a continuous state space, with normally distributed observations — a Wiener process. The choice of the mean and variance in the normal distribution is related to an approximation of the non-zero root h_0 to the equation $\ln \varphi_\lambda(h) = 0$. An analytic expression of h_0, chosen in such a way that it gives the correct values of h_0 when λ equals λ_0 and λ_1, was proposed in Section II.2.

The validity of these approximations can be studied on two levels. First, the approximation of the d.f. of N in a non-truncated SPRT is examined. This is accomplished by studying the first central moments of the distribution of N. On the second level, approximations to the characteristics of the decisive sampling number in truncated SPRT is studied.

The d.f. of N in a non-truncated SPRT can be evaluated in two steps. In the first step, approximations are made of the discrete state space to a continuous state space. In the second step, the approximation of the discrete time process to a continuous time process is evaluated.

IV.2 Approximation of the Discrete State Space to a Continuous State Space

When estimating the error made by approximating the discrete state space — where the observations are Poissons distributed — to a continuous state space — where the observations are normally distributed — the d.f. of N is compared with the distribution function of the decisive sampling time T, defined as follows. Let X(t) be a Poisson process, i.e., the probability mass function of X(t) is given by

$$f_t(x;\lambda) = \frac{(\lambda t)^x \cdot e^{-\lambda t}}{x!} \quad \text{for} \quad x = 0,1,2 \ldots . \quad \text{and} \; t > 0$$

$$f_0(0;\lambda) = 1 \qquad \text{(IV.2.1)}$$

$$f_0(x;\lambda) = 0 \quad \text{for} \quad x \neq 0 .$$

Given two risk levels α and β, such that $\alpha + \beta < 1$ we now want to decide between $H_0: \lambda = \lambda_0$ and $H_1: \lambda = \lambda_1$. The optimal test procedure (optimal in the sense that it minimizes $E(T;\lambda_0)$ and $E(T;\lambda_1)$) is then given by Wald's SPRT, with the test statistic

$$S(t) = X(t) \ln (\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0)\, t \qquad \text{(IV.2.2)}$$

for $t > 0$ and $S(0) = 0$. Dvoretzky et al. (1953) studied this situation and gave exact expressions of the OC-function and the moment-generating function of T. The adjustments of their derivation which are necessary in order to make it formally valid are given in Hall (1970) and Zacks (1971). Let r be the start value of the process

$$R(t) = \frac{S(t)}{\ln(\lambda_1/\lambda_0)} = X(t) - \frac{(\lambda_1 - \lambda_0)}{\ln(\lambda_1/\lambda_0)} \cdot t , \qquad \text{(IV.2.3)}$$

that is, $r = R(0)$ and a and b are the boundaries to the continuation region of the test depending on the risks α and β. Two constants, K and J, are defined as

$$K = \frac{b}{\ln(\lambda_1/\lambda_0)} \; ; \; J = \frac{a}{\ln(\lambda_1/\lambda_0)} . \qquad \text{(IV.2.4)}$$

Let $n(r)$ be the integer such that

$$J-r-1 \leq n(r) < J-r \ . \qquad \text{(IV.2.5)}$$

For $K \leq r < J$, the moment-generating function of T is then given by (cf. Dvoretzky et al. (1953)):

$$M_\lambda(u;r) = \left(\frac{\lambda}{\lambda-u}\right)^{[n(r)+1]} + C(u)e^{-[r(\lambda-u)/c]} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} \left\{(J-r-i)\frac{\lambda}{c} e^{-[(\lambda-u)/c]}\right\}^i$$

$$- \frac{\lambda u}{(\lambda-u)^2} e^{-[(\lambda-u)(J-r-1)/c]} \sum_{i=0}^{n(r)-1} \left[\frac{\lambda}{\lambda-u} e^{-(\lambda-u)/c}\right]^i \sum_{j=0}^{i} \left[(J-r-i)\frac{\lambda-u}{c}\right]^j \qquad \text{(IV.2.6)}$$

where $C(u)$ is determined so that $M_\lambda(u;K) \equiv 1$.

Expressions of the first three central moments of T can be derived from this expression of the moment-generating function of T. This derivation is presented in Appendix C. Then, values of the three first central moments of T were calculated for different values of α, β, λ_0, λ_1 and λ.

When the values of the moments were calculated, the numerical calculations were found to be very sensitive to the calculation method and the degree of precision. The calculations were made separately for each sum and in increasing order for the indexes. A computation procedure where all terms for each value of the index i are added probably requires a lower degree of precision for the values at each stage, but this has not been tested. Instead, different degrees of precision for the values at each stage were used. The calculations were made on the IBM 360 and later on the IBM 370 and were made in single — giving six to seven relevant figures — double — about seventeen relevant figures — and 'extended' precision — about thirty-five relevant figures. Table IV.2.1 shows values of the three central moments of T calculated using varying degrees of precision for the case where $\alpha = \beta = 0.05$, $\lambda_0 = 5.0$, $\lambda_1 = 5.5$, 6.25 and 7.50 and for values of λ equal to λ_0, $(\lambda_1+\lambda_0)/2$ and λ_1. If two adjacent levels of precision yield the same or nearly the same result, the numerical error in the

calculated values is small. This table implies that when $\lambda_0 = 5.0$ and $\lambda_1 = 5.5$, it is not certain whether 'extended' precision will lead to small numerical errors for any of the moments. For $\lambda_0 = 5.0$ and $\lambda_1 = 6.25$, 'extended' precision certainly gives small numerical errors for E(T) and V(T), but it is not clear whether this is valid for $E(T-\mu_T)^3$. When $\lambda_0 = 5.0$ and $\lambda_1 = 7.5$, 'extended' precision leads to small numerical errors for all three moments. A study of other cases shows that the necessary degree of precision in the calculations depends on the risk levels α and β and the ratio λ_1/λ_0. The values studied were $\alpha = \beta = 0.01$, $\alpha = \beta = 0.05$ and $\alpha = 0.01$, $\beta = 0.05$, $\lambda_0 = 1.0$, 5.0, 10,0 and 15.0 and the ratio $\lambda_1/\lambda_0 = 1.10$, 1.25 and 1.50.

For these cases, single precision does not lead to small numerial errors in any of the three moments of T. Double precision gave accurate values for all three moments of T when $\alpha = \beta = 0.05$ for ratios $\lambda_1/\lambda_0 = 1.25$ and 1.50 and for other risk levels only when $\lambda_1/\lambda_0 = 1.50$. In all of the other cases, it did not give accurate values for any of the moments. Finally, 'extended' precision gave reliable results for all risk levels when the ratio λ_1/λ_0 was larger than or equal to 1.25.

Calculations of the moments of N are also sensitive to the degree of precision. For $\alpha = \beta = 0.05$, double precision gave accurate values, while 'extended' precision had to be used for the other risk levels. In the expression of the moments of T and N, λ_0 can be treated as a scale factor. Appendix F shows the values of the moments of N and T when λ_0 equals 1.0. When the risk levels are $\alpha = \beta = 0.05$ also values for λ_0 equal to 5.0 and 10.0 are presented, in order to make the comparison to values estimated from simulations simple.

The stochastic process given by S(t) will have no excess over b at termination if H_0 is accepted. However, S(t) will have an excess over a if the test is terminated with acceptance of H_1. When the state space of this process is approxi-

mated by means of a continuous state space, the excess of S(t) over a at termination is neglected. The approximation is chosen so that the log likelihood ratio Z will have the same variance as in the Poisson case, except that the expected value of Z is changed. The true expected value of Z

$$E(Z) = \lambda \ln(\lambda_1/\lambda_0) - (\lambda_1-\lambda_0) \qquad \text{(IV.2.7)}$$

is changed to

$$\frac{(\lambda_1+\lambda_0-2\lambda)}{2(\lambda_1-\lambda_0)} \sqrt{\lambda \ln(\lambda_1/\lambda_0)} \; . \qquad \text{(IV.2.8)}$$

Figure II.2.1 shows that for values of λ less than λ_2 — where $\lambda_2 \simeq \lambda_0 + \frac{1}{4}(\lambda_1-\lambda_0)$ — the approximation uses a higher mean and for values of λ larger than λ_2, the approximation uses a lower mean. This is aimed at making the approximation more valid. It was also proposed that the true mean of Z should be used in the approximation for values of λ near $\lambda' = (\lambda_1-\lambda_0)/\ln(\lambda_1/\lambda_0)$; cf. Section II.2. The validity of the approximation depends on the approximation of the root h_0 to the equation $\ln\varphi_\lambda(h) = 0$, where $\varphi_\lambda(h)$ is the moment-generating function of Z.

A comparison of moments of T with moments of N for $\lambda = \lambda_0$ and λ_1 will indicate the effect of neglecting the excess of S(t) over a at termination of the test. The same comparison for $\lambda = (\lambda_1+\lambda_0)/2$ will also indicate the effect of errors in h_0.

Table IV.2.1 contains values of the first central moments of T and of N — in the latter case, values calculated from the approximation to the characteristic functions of N as well as estimates from simulations (see Section IV.3) — when $\alpha = \beta = 0.05$. Neglection of the excess of S_N over a at termination of the test leads to an underestimate of the central moments. As can be predicted the relative underestimate increases with the ratio λ_1/λ_0. Table IV.2.2 presents the values of the coefficients of variation, skewness and excess for N. Table IV.2.3 lists the values of the coefficients of variation and excess for T. A comparison of these two tables indicates that, for $\lambda = \lambda_0$ and λ_1, the

approximation used in this study yields accurate values of the coefficients of variation and skewness when testing a continuous Poisson process. The differences are so small that they could be explained by numerical approximations and errors. The approximation to the average sampling number will underestimate the average sampling time when λ equals λ_1. When λ_1/λ_0 equals 1.5 this underestimate is still less than 5 per cent.

Moments of N for $\lambda = (\lambda_1+\lambda_0)/2$ were calculated in order to estimate the effect of an error in the root h_0 to the equation $\ln \varphi_\lambda(h) = 0$. In this case the approximation (IV.2.8) of the mean of the log likelihood ratio Z could not be used because it equals zero. Instead, the true value, given by (IV.2.7), was used. The 'double' precision on the IBM 360 was too low to give accurate values for more than E(N). Table IV.2.4 shows the expected values of N and T. This table indicates that E(N) underestimates E(T) somewhat and that the underestimate increases with the ratio λ_1/λ_0. When λ_1/λ_0 equals 1.5, this underestimate is less than 5 per cent. Use of 'extended' precision in the calculations neither gave accurate values of the other central moments.

IV.3 Error Made by Approximating the Discrete Time Process to a Continuous Time Process

In order to study the error when approximating the discrete time process to a continuous time process, simulations of SPRT for testing the parameter in a Poisson distribution were made and the first four central moments of the decisive sampling number, N, were estimated from the simulations. Simulations were performed for risk levels $\alpha = \beta = 0.05$, values of $\lambda_0 = 5.0$ and 10.0 and the ratio λ_1/λ_0 equal to 1.1, 1.25 and 1.50. During the simulations, the SPRT had to be truncated so as not to obtain a sample with an extremely high value of N. A rule proposed by Gosh (1970) was used first. He stated that if the log likelihood ratios, Z_i, are independent and identically distributed, a useful working rule to assure that $P(N \le N_0|\lambda) \approx 1$ is to take $N_0 = 3 \cdot \sup_\lambda \tilde{E}(N;\lambda)$,

where $\tilde{E}(N;\lambda)$ is Wald's approximation to the average sampling number, described in Section I.3. As a less strict rule, the truncation point N_0 could be chosen as $3 \cdot \max[\tilde{E}(N;\lambda_0), \tilde{E}(N;\lambda_1)]$. Wald (1947) and Gosh (1970) have studied upper bounds for the risks of error in truncated SPRT with different truncation points N_0 and showed that the change in the risk levels will be small using these two rules, which are identical from this aspect. Gosh (1970, p. 143) also assigned a lower bound to $P(N < N_0|\mu_0)$ and $P(N < N_0|\mu_1)$ of the UME SPRT of $H_0: \mu < \mu_0$ against $H_1: \mu > \mu_1$ in a normal distribution with known variance for different risk levels. When the simulations were performed it was found that the truncation point determined by either of these two rules was too low to give a good estimate of the coefficients of skewness and excess for the decisive sampling number N. Two methods for shoosing the truncation point were studied for two cases, namely $\lambda_0 = 5.0$, $\lambda_1 = 6.25$ and $\lambda_0 = 5.0$, $\lambda_1 = 7.5$. The first method requires that the same truncation point N_0 should be used for all values of λ and that N_0 be chosen as

$$N_0 = k \cdot \sup[\tilde{E}(N;\lambda_0), \tilde{E}(N;\lambda_1)] . \qquad \text{(IV.3.1)}$$

This rule is designated G(k). The second method requires that a special truncation point be chosen for each value of λ, chosen as

$$N_0 = k \cdot \tilde{E}(N;\lambda) . \qquad \text{(IV.3.2)}$$

This rule is designated S(k). Sampling values of the moments and the coefficients of variation, skewness and excess for the two cases studied, (k = 3,4,5,6), are listed in Tables IV.3.1-2. It can be seen that application of the general rule G(k) given by (IV.3.1) with k = 6 will result in underestimates of all parameters. That these underestimates are not quite negligible can be seen from the case $\lambda_0 = 5.0$, $\lambda_1 = 7.5$, $\lambda = 7.5$ where the coefficient of excess is still underestimated by more than 10 per cent. For $\lambda = (\lambda_1+\lambda_0)/2$, the underestimates of the coefficients of skewness and excess are serious using this rule and their size has been estimated at 40-50 per cent. This means that in order to obtain

reliable estimates of expected value and coefficients of variation, skewness and excess for the decisive sampling number in SPRT the truncation point has to be chosen as least as large as $6 \cdot E(N;\lambda)$. This rule was adhered to the simulations, except for the case $\lambda_0 = 1.0$ and $\lambda_1 = 1.25$ and 1.50 (where Gosh's rule was used and for which only the expected value and the coefficient of variation are recorded).

The sample variation has to be taken into account when studying the sample estimates of the characteristics of N. Sample estimates of the variance for the estimate $E(N;\lambda)$ will be given by

$$\hat{V}\{\hat{E}(N;\lambda)\} = \hat{V}(N;\lambda)/n$$

where n is the number of simulations. For the estimate of the coefficients of variation, $\hat{\omega}$, skewness, $\hat{\gamma}_1$, and excess, $\hat{\gamma}_2$, results given by Cramér (1945) can be used. They are as follows when the following notations are used:

$$E(N;\lambda) = \mu$$
$$V(N;\lambda) = \mu_2 \qquad \sqrt{V(N;\lambda)} = \sigma \qquad \text{(IV.3.3)}$$
$$E\{(N-\mu)^k;\lambda\} = \mu_k$$

$$E(\hat{\omega}) \cong \frac{\sigma}{\mu} \quad V(\hat{\omega}) \cong \frac{\mu^2(\mu_4 - \mu_2^2) - 4\mu\mu_2\mu_3 + 4\mu_2^3}{4\mu^4\mu_2 n}$$

$$E(\hat{\gamma}) \cong \gamma_1 \quad V(\hat{\gamma}_1) \cong \frac{4\mu_2^2\mu_6 - 12\mu_2\mu_3\mu_5 - 24\mu_2^3\mu_4 + 9\mu_3^2\mu_4 + 35\mu_2^2\mu_3^2 + 36\mu_2^5}{4\mu_2^5\mu n}$$

$$E(\hat{\gamma}_2) \cong \gamma_2 \quad V(\hat{\gamma}_2) \cong \frac{\mu_2^2\mu_8 - 4\mu_2\mu_4\mu_6 - 8\mu_2^2\mu_3\mu_5 + 4\mu_4^4 - \mu_2^2\mu_4^2 + 16\mu_2\mu_3^2\mu_4 + 16\mu_2^3\mu_3^2}{\mu_2^6 n}$$

Terms of order n^{-1} for the expected value and of order $n^{-3/2}$ for the variance are neglected in the approximations. Cramér also shows that $\hat{\omega}$, $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are asymptotically normally distributed for large values of n. Table IV.3.3 contains estimates of the standard deviation of $\hat{E}(N;\lambda)$ and $\hat{\omega}$ using sample

estimates of the central moments for the cases studied. The tables indicate that the relative error in the sample estimate of the expected value of N varies between one and two per cent except for the case when $\lambda_0 = 5.0$ and $\lambda_1 = 6.25$, where it is smaller than half a per cent. For the sample estimate of the coefficient of variation of N, the relative error varies between one and three per cent except for the same case, where it is about half a per cent. The relative error in the sample estimates of the coefficients of skewness and excess have been estimated for the case $\lambda_0 = 10.0$, $\lambda_1 = 15.0$ and $\lambda = 15.0$. The following values were obtained for this case:

$\hat{\gamma}_1 = 0.621$ $\quad \hat{\sigma}(\hat{\gamma}_1) = 0.089$ $\quad \hat{\sigma}(\hat{\gamma}_1)/\hat{\gamma}_1 = 0.055$

$\hat{\gamma}_2 = 6.64$ $\quad \hat{\sigma}(\hat{\gamma}_2) = 0.574$ $\quad \hat{\sigma}(\hat{\gamma}_2)/\hat{\gamma}_2 = 0.086$

The sample estimates of the average sampling number and the coefficients of variation, skewness and excess of N are presented in Table IV.3.4. The estimated moments of N are presented in Appendix F. The effect of approximating the discrete time process to a continuous time process is studied on the basis of the ratio between characteristics of the decisive sampling number, N, estimated from simulations and characteristics of the decisive sampling time, T.

These ratios for the average sampling number and the coefficients of variation and skewness are shown in Table IV.3.5. The ratio between sample estimates and values calculated from the approximation to the ch.f. of N are given for the coefficient of excess.

According to Table IV.3.5, when $\alpha = \beta = 0.05$, approximation of the discrete time Poisson process to a continuous time Poisson process and application of the characteristics of the decisive sampling time T will lead to

i) an underestimate of the average sampling number E(N). This underestimate depends on λ and will be about the same for $\lambda = \lambda_0$ and λ_1 and larger for $\lambda = (\lambda_1+\lambda_0)/2$. The underestimate increases with λ_0 and the ratio λ_1/λ_0.

ii) an overestimate of the coefficient of variation of N. In most cases the differences are so small that c.v. of T falls within the 95 per cent confidence intervals estimated from the simulations.

iii) an overestimate of the coefficient of skewness. When λ equals λ_0 and λ_1 the differences are mostly so small that they ought to fall within the 95 per cent confidence intervals of the sample estimates.

The results above also hold when the continuous time process with normally distributed state space as an approximation is used to estimate the characteristics of the decisive sampling time in a SPRT. When estimating the ASN-function it should be kept in mind that this function will be underestimated except when λ_0 and λ_1/λ_0 are small. When the approximation is used to estimate the coefficients of variation, skewness and excess, overestimates will result. The overestimates are moderate for the first two coefficients but, with respect to the coefficient of excess, it is not clear that it always will be moderate.

IV.4 Evaluating Approximations of Characteristics of Truncated SPRT

One of the main purposes of this paper has been to study truncated tests. Expressions of the OC and ASN-functions have been derived. We now turn to the ASN-function and make an attempt to determine the accuracy of the approximation.

One-sided and two-sided truncated SPRT of the hypothesis $H_0: \lambda \leq \lambda_0$ against $H_1: \lambda \geq \lambda_1$ have been simulated for the risk levels $\alpha = \beta = 0.05$, $\lambda_0 = 5.0$ and 10.0 and $\lambda_1/\lambda_0 = 1.1$, 1.25 and 1.50. Table IV.4.1 contains sample estimates of the average sampling number, $\hat{E}(N)$, sample estimates of the standard deviation of $\hat{E}(N)$ and values calculated from the approximate expressions of the ASN-function given in Section III.2. Ratios between sample estimates $\hat{E}(N)$ and calculated values of $E(N)$ for $\lambda = \lambda_0$ and λ_1 are presented in Table IV.4.2. This table also contains a comparison of two-sided truncated SPRT with different truncation points, N_0. Use of the approximation to estimate the average sampling

number will, as for the non-truncated SPRT, lead to an underestimate that increases with λ_0 and the ratio λ_1/λ_0. When λ equals λ_0, the underestimate will be smaller than in the SPRT and when λ equals λ_1 it can be large. When the truncation point, N_0, in two-sided SPRT is chosen as low as the fixed-sample size, the underestimate is considerably lower than in the SPRT.

In the next chapter the actual risk levels for the tests will also be considered and it can then be shown that the higher values of the estimate $\hat{E}(N)$ can, at least to some extent, be explained by the fact that the actual risk levels were lower than the intended ones.

Table IV.2.1 Values of the first four central moments of the decisive sampling number, N, and the first three central moments of the decisive sampling time, T. The values for the decisive sampling number, N, are calculated from an approximation of the characteristic function of N, denoted "approx.", and estimated from simulations, denoted "simulation".

λ_0	λ_1	λ	$E(T)=\mu_T$	$E(N)=\mu_N$ approx.	$E(N)=\mu_N$ simulation	$V(T)$	$V(N)$ approx.	$V(N)$ simulation	$E(T-\mu_T)^3$	$E(N-\mu_N)^3$ approx.	$E(N-\mu_N)^3$ simulation	$E(N-\mu_N)^4$ approx.	$E(N-\mu_N)^4$ simulation
5.00	5.50	5.00		113.01	117.01		6433	6062		$0.97 \cdot 10^6$	$0.69 \cdot 10^6$	$3.52 \cdot 10^8$	$2.07 \cdot 10^8$
		5.50		109.48	114.41		6433	6906		$0.97 \cdot 10^6$	$1.01 \cdot 10^6$	$3.52 \cdot 10^8$	$3.76 \cdot 10^8$
	6.25	5.00	19.87	19.73	22.40	192	188	234	5040	4859	6748	$3.00 \cdot 10^5$	$4.50 \cdot 10^5$
		6.25	18.81	18.32	21.57	194	188	234	5074	4859	6440	$3.00 \cdot 10^5$	$4.26 \cdot 10^5$
	7.50	5.00	5.67	5.61	7.27	14.72	14.19	21.73	108	101	180	1717	3683
		7.50	5.14	4.90	6.70	15.06	14.20	23.03	109	101	207	1718	4337
10.0	11.0	10.0		56.51	60.93		1608	1759		$1.22 \cdot 10^5$	$1.27 \cdot 10^5$	$2.20 \cdot 10^7$	$2.16 \cdot 10^7$
		11.0		54.74	60.34		1608	1760		$1.22 \cdot 10^5$	$1.39 \cdot 10^5$	$2.20 \cdot 10^7$	$2.57 \cdot 10^7$
10.0	12.5	10.0	9.93	9.87	11.71	47.91	46.92	59.34	630	607	784	$1.88 \cdot 10^4$	$2.44 \cdot 10^4$
		12.5	9.41	9.16	11.53	48.50	46.92	63.44	634	607	843	$1.88 \cdot 10^4$	$2.68 \cdot 10^4$
10.0	15.0	10.0	2.84	2.80	3.95	3.68	3.55	5.91	13.50	12.64	26.83	107	295
		15.0	2.57	2.45	3.65	3.77	3.55	5.69	13.62	12.64	21.88	107	215

Table IV.2.2 Coefficients of variation, skewness and excess of the decisive sampling number, N, calculated from an approximate expression of the characteristic function of N.

$\Delta = \frac{\lambda_1}{\lambda_0} - 1$		risk levels		coefficient of		
Δ	λ	α	β	variation	skewness	excess
0.10	λ_0	0.05	0.05	0.71	1.89	5.51
	λ_1			0.73	1.89	5.51
0.25	λ_0			0.69	1.89	5.51
	λ_1			0.75	1.89	5.51
0.50	λ_0			0.67	1.89	5.51
	λ_1			0.77	1.89	5.51
0.10	λ_0	0.01	0.01	0.62	1.78	4.98
	λ_1			0.65	1.78	4.98
0.25	λ_0			0.61	1.78	5.00
	λ_1			0.63	1.78	5.00
0.50	λ_0			0.59	1.78	5.03
	λ_1			0.68	1.78	5.03
0.10	λ_0	0.01	0.05	0.76	2.10	6.78
	λ_1			0.63	1.67	4.41
0.25	λ_0			0.74	2.10	6.81
	λ_1			0.64	1.67	4.45
0.50	λ_0			0.72	2.10	6.86
	λ_1			0.65	1.68	4.50

Table IV.2.3 Coefficient of variation and skewness of the decisive sampling time, T, calculated from an exact expression of the moment-generating functions of T.

$\Delta = \frac{\lambda_1}{\lambda_0} - 1$		risk levels		Coefficient of	
Δ	λ	α	β	variation	skewness
0.10	λ_0	0.05	0.05		
	λ_1				
0.25	λ_0			0.70	1.90
	λ_1			0.74	1.88
0.50	λ_0			0.68	1.91
	λ_1			0.76	1.87
0.10	λ_0	0.01	0.01		
	λ_1				
0.25	λ_0			0.61	1.79
	λ_1			0.65	1.77
0.50	λ_0			0.59	1.79
	λ_1			0.67	1.76
0.10	λ_0	0.01	0.05		
	λ_1				
0.25	λ_0			0.75	2.11
	λ_1			0.63	1.66
0.50	λ_0			0.72	2.12
	λ_1			0.65	1.67

Table IV.2.4 Average sampling number derived from an approximate expression of the characteristic function of N compared with average sampling time when $\lambda = (\lambda_1+\lambda_0)/2$.

			Expected values					
		$\lambda =$	$\alpha = \beta = 0.05$		$\alpha = \beta = 0.01$		$\alpha = 0.01$ $\beta = 0.05$	
λ_0	λ_1	$(\lambda_1+\lambda_0)/2$	E(N)	E(T)	E(N)	E(T)	E(N)	E(T)
1.00	1.10	1.05	908.94		2213.74		1443.27	
1.00	1.25	1.125	154.62	159.05	376.03	383.36	244.69	249.50
1.00	1.50	1.25	42.05	44.47	101.96	106.06	66.97	69.71
5.00	5.50	5.25	181.79		442.75		288.65	
5.00	6.25	5.625	30.92	31.81	75.21	76.67	48.94	49.90
5.00	7.50	6.25	8.41	8.89	20.39	21.21	13.39	13.94
10.00	11.00	10.50	90.88		221.28		143.10	
10.00	12.50	11.25	15.46	15.91	37.60	38.34	24.47	24.95
10.00	15.00	12.50	4.21	4.45	10.20	10.61	6.70	6.97
15.00	16.50	15.75	60.59		147.52		95.40	
15.00	18.75	16.875	10.31	10.60	25.07	25.26	16.31	16.63

Table IV.3.1 Values of sampling moments of N calculated from simulated SPRT, when different truncation points, N_0, are used. The rules for determining truncation points are called G(k) and S(k), where G = general denotes use of the same truncation point for all values of λ, that is $N_0 \simeq k \cdot \max [\tilde{E}(N;\lambda_0), \tilde{E}(N;\lambda_1)]$ and S = specific denotes use of a specific truncation point for each value of λ, that is $N \simeq k \cdot \tilde{E}(N;\lambda)$. Number of simulations 30,000.

λ_0	λ_1	λ	Method for determining N_0	Truncation point N_0	Sample estimates $\hat{E}(N) = \hat{\mu}_N$	$\hat{V}(N)$	$\hat{E}(N-\hat{\mu}_N)^3$	$\hat{E}(N-\hat{\mu}_N)^4$	Coefficient of variation	skewness	excess
5.000	6.250	5.000	G(3), S(3)	60	21.93	185	2921	$1.28 \cdot 10^5$	0.620	1.161	0.740
			G(4), S(4)	79	22.27	216	4897	$2.59 \cdot 10^5$	0.660	1.543	2.551
			G(5), S(5)	99	22.37	229	6054	$3.63 \cdot 10^5$	0.676	1.747	3.922
			G(6), S(6)	119	22.39	233	6545	$4.20 \cdot 10^5$	0.682	1.840	4.736
			no truncation		22.40	234	6748	$4.50 \cdot 10^5$	0.683	1.885	5.218
5.000	6.250	5.625	G(3)	60	32.56	339	1689	$1.93 \cdot 10^5$	0.565	0.271	
			G(4)	79	35.14	518	7924	$6.05 \cdot 10^5$	0.648	0.672	
			G(5)	99	36.57	666	17102	$1.37 \cdot 10^6$	0.706	0.995	0.089
			G(6)	119	37.30	770	26654	$2.37 \cdot 10^6$	0.744	1.247	0.997
			no truncation		38.10	935	52360	$6.68 \cdot 10^6$	0.803	1.831	4.641
5.000	6.250	5.625	S(3)	93	36.24	627	14219	$1.11 \cdot 10^6$	0.691	0.906	
			S(4)	124	37.42	790	28849	$2.63 \cdot 10^6$	0.751	1.299	1.214
			S(5)	155	37.89	878	40447	$4.26 \cdot 10^6$	0.782	1.555	2.526
			S(6)	186	38.08	929	50736	$6.26 \cdot 10^6$	0.800	1.792	4.253
			no truncation		38.10	935	52360	$6.68 \cdot 10^6$	0.803	1.831	4.641
5.000	6.250	6.250	G(3)	60	21.14	189	2936	$1.32 \cdot 10^5$	0.650	1.130	0.695
			G(4)	79	21.46	218	4793	$2.57 \cdot 10^5$	0.688	1.489	2.408
			G(5)	99	21.54	229	5858	$3.54 \cdot 10^5$	0.703	1.690	3.750
			G(6)	119	21.56	233	6297	$4.05 \cdot 10^5$	0.708	1.771	4.460
			S(3)	55	20.97	177	2364	$1.01 \cdot 10^5$	0.634	1.004	0.224
			S(4)	74	21.41	213	4381	$2.25 \cdot 10^5$	0.682	1.409	1.959
			S(5)	92	21.53	227	5591	$3.27 \cdot 10^5$	0.700	1.635	3.346
			S(6)	110	21.56	232	6169	$3.89 \cdot 10^5$	0.706	1.746	4.227
			no truncation		21.57	234	6440	$4.26 \cdot 10^5$	0.709	1.799	4.780

Table IV.3.1 Estimated probability of obtaining values of N larger than or equal to the truncation point N_0 and ratio between values for truncated tests and non-truncated tests of the coefficients of variation, skewness and excess. Numbers of simulations 30,000.

λ_0	λ_1	λ	Method for determining N_0	Estimated $P(N \geq N_0)$	Ratio between values for truncated tests and non-truncated tests of coefficient of variation	skewness	excess
5.000	6.250	5.000	G(3), S(3)	0.0320	0.91	0.62	0.14
			G(4), S(4)	0.0095	0.97	0.82	0.49
			G(5), S(5)	0.0022	0.99	0.93	0.75
			G(6), S(6)	0.0006	1.00	0.98	0.91
5.000	6.250	5.625	G(3)	0.1847	0.70	0.15	
			G(4)	0.1003	0.81	0.37	
			G(5)	0.0512	0.88	0.54	0.02
			G(6)	0.0263	0.93	0.68	0.21
			S(3)	0.0621	0.86	0.49	
			S(4)	0.0217	0.94	0.71	0.26
			S(5)	0.0076	0.97	0.85	0.54
			S(6)	0.0025	1.00	0.98	0.92
5.000	6.250	6.250	G(3)	0.0289	0.92	0.63	0.15
			G(4)	0.0089	0.97	0.83	0.50
			G(5)	0.0018	0.99	0.94	0.78
			G(6)	0.0005	1.00	0.98	0.93
			S(3)	0.0409	0.89	0.56	0.05
			S(4)	0.0121	0.96	0.78	0.41
			S(5)	0.0030	0.99	0.91	0.70
			S(6)	0.0009	1.00	0.97	0.88

Table IV.3.2 Values of sampling moments of N calculated from simulated SPRT, when different truncation points, N_0, are used. The rules for determining truncation points are called G(k) and S(k), where G = general denotes use of the same truncation point for all values of λ, that is $N_0 \simeq k \cdot \max [\tilde{E}(N; \lambda_0), \tilde{E}(N; \lambda_1)]$ and S = specific denotes use of a specific truncation point for each value of λ, that is $N_0 \simeq k \cdot \tilde{E}(N; \lambda)$. Number of simulations 2,000.

λ_0	λ_1	λ	Method for determining N_0	Truncation point N_0	Sample estimates $\hat{E}(N) = \hat{\mu}_N$	$\hat{V}(N)$	$\hat{E}(N - \hat{\mu}_N)^3$	$\hat{E}(N) - \hat{\mu}_N)^4$	Coefficient of variation	skewness	excess
5.000	7.500	5.000	G(3), S(3)	17	7.07	16.38	60.47	810	0.572	0.912	0.019
			G(4), S(4)	23	7.21	19.79	116.94	1848	0.617	1.328	1.719
			G(5), S(5)	28	7.25	21.17	151.69	2717	0.635	1.557	3.062
			G(6), S(6)	34	7.26	21.73	170.11	3305	0.642	1.679	3.999
			no truncation		7.27	21.96	179.77	3683	0.645	1.747	4.637
5.000	7.500	6.250	G(3)	17	9.99	28.12	13.98	1232	0.531	0.094	
			G(4)	23	9.50	28.36	104.05	2243	0.561	0.689	
			G(5)	28	11.57	61.74	381.23	9435	0.679	0.786	
			G(6)	34	11.89	73.91	664.81	17726	0.723	1.046	0.245
			no truncation		12.23	94.26	1525.90	54736	0.794	1.667	3.161
5.000	7.500	6.250	S(3)	25	11.31	53.95	250.65	6231	0.649	0.633	
			S(4)	34	11.89	73.91	664.81	17726	0.723	1.046	0.245
			S(5)	42	12.10	84.65	1027.8	31248	0.760	1.320	1.361
			S(6)	51	12.19	90.96	1324.0	45408	0.782	1.526	2.488
			no truncation		12.23	94.26	1525.90	54736	0.794	1.667	3.161
5.000	7.500	7.500	G(3)	17	6.50	16.95	72.90	948	0.633	1.045	0.300
			G(4)	23	6.65	20.86	140.66	2234	0.687	1.476	2.134
			G(5)	28	6.69	22.24	176.05	3136	0.705	1.679	3.340
			G(6)	34	6.70	22.76	191.94	3651	0.711	1.774	4.079
			S(3)	15	6.40	15.02	50.10	620	0.606	0.861	
			S(4)	20	6.60	19.30	109.21	1575	0.666	1.288	1.228
			S(5)	25	6.67	21.65	159.86	2698	0.698	1.587	2.756
			S(6)	30	6.69	22.46	183.18	3355	0.708	1.721	3.651
			no truncation		6.70	23.03	207.37	4337	0.716	1.876	5.177

Table IV.3.2 Estimated probability of obtaining values of N larger than or equal to the truncation point N_0 and ratio between values for truncated tests and non-truncated tests of the coefficients of variation, skewness and excess. Numbers of simulations 2,000.

λ_0	λ_1	λ	Method for determining N_0	Estimated $P(N \geq N_0)$	Ratio between values for truncated tests and non-truncated tests of coefficient of variation	skewness	excess
5.000	7.500	5.000	G(3), S(3)	0.047	0.89	0.52	0.00
			G(4), S(4)	0.014	0.96	0.76	0.37
			G(5), S(5)	0.004	0.98	0.89	0.66
			G(6), S(6)	0.001	1.00	0.96	0.86
5.000	7.500	6.250	G(3)	0.251	0.67	0.06	
			G(4)	0.135	0.71	0.41	
			G(5)	0.078	0.86	0.47	
			G(6)	0.040	0.91	0.63	0.08
			S(3)	0.111	0.82	0.38	
			S(4)	0.040	0.91	0.63	0.08
			S(5)	0.017	0.96	0.79	0.43
			S(6)	0.007	0.98	0.92	0.79
5.000	7.500	7.500	G(3)	0.045	0.88	0.56	0.06
			G(4)	0.014	0.96	0.79	0.41
			G(5)	0.003	0.98	0.89	0.65
			G(6)	0.001	0.99	0.95	0.79
			S(3)	0.072	0.85	0.46	
			S(4)	0.027	0.93	0.69	0.23
			S(5)	0.011	0.97	0.85	0.53
			S(6)	0.003	0.99	0.92	0.71

Table IV.3.3 Sample estimates of the standard deviations of $\hat{E}(N;\lambda)$ and $\hat{\omega}$, the coefficient of variation for N.
Risk levels $\alpha = \beta = 0.05$.

λ_0	λ_1	λ	Number of simulations	Sample estimates $\hat{D}(\hat{E})$	$\frac{\hat{D}(\hat{E})}{\hat{E}}$	$\hat{D}(\hat{\omega})$	$\frac{\hat{D}(\hat{\omega})}{\hat{\omega}}$
1.00	1.25	1.00	1500	1.76	0.017	0.0134	0.020
		1.125		3.33	0.020	0.0134	0.017
		1.25		1.86	0.019	0.0142	0.019
1.00	1.50	1.00	3000	0.371	0.012	0.0107	0.016
		1.25		0.733	0.015	0.0104	0.012
		1.50		0.369	0.013	0.0125	0.017
5.00	5.50	5.00	1500	2.01	0.017	0.0136	0.020
		5.25		4.16	0.021	0.0156	0.019
		5.50		2.15	0.019	0.0185	0.025
5.00	6.25	5.00	30000	0.088	0.004	0.0039	0.006
		5.625		0.177	0.005	0.0042	0.005
		6.25		0.088	0.004	0.0040	0.006
5.00	7.50	5.00	2000	0.104	0.014	0.0141	0.022
		6.25		0.217	0.018	0.0137	0.017
		7.50		0.107	0.016	0.0157	0.022
10.00	11.00	10.00	1500	1.08	0.018	0.0157	0.023
		10.50		2.13	0.021	0.0173	0.022
		11.00		1.08	0.018	0.0179	0.026
10.00	12.50	10.00	3000	0.141	0.012	0.0106	0.016
		11.25		0.292	0.014	0.0112	0.014
		12.50		0.145	0.013	0.0109	0.016
10.00	15.00	10.00	3000	0.044	0.011	0.0118	0.019
		12.50		0.097	0.014	0.0089	0.012
		15.00		0.044	0.012	0.0106	0.016

Table IV.3.4 Sample estimates of the average sampling number and the coefficients of variation, skewness and excess for N, estimated from simulations of SPRT with risk levels $\alpha = \beta = 0.05$.

λ_0	λ_1	λ	Number of simulations	Estimates of: Average sampling number $\hat{E}(N)$	Coefficient of variation $\hat{\omega}$	Coefficient of skewness $\hat{\gamma}_1$	Coefficent of excess $\hat{\gamma}_2$
1.00	1.25	1.00	1500	103.81	0.66		
		1.125		163.49	0.79		
		1.25		96.69	0.74		
1.00	1.50	1.00	3000	31.11	0.65		
		1.25		47.86	0.84		
		1.50		27.67	0.73		
5.00	5.50	5.00	1500	117.01	0.67	1.46	2.63
		5.25		198.42	0.81	1.64	2.91
		5.50		114.41	0.73	1.76	4.88
5.00	6.25	5.00	30000	22.40	0.68	1.89	5.22
		5.625		38.10	0.80	1.83	4.64
		6.25		21.57	0.71	1.80	4.78
5.00	7.50	5.00	2000	7.27	0.64	1.77	4.80
		6.25		12.23	0.79	1.67	3.16
		7.50		6.70	0.72	1.88	5.18
10.00	11.00	10.00	1500	60.93	0.69	1.72	3.98
		10.50		102.87	0.80	1.69	3.65
		11.00		60.34	0.70	1.88	5.30
10.00	12.50	10.00	3000	11.71	0.66	1.72	3.92
		11.25		20.60	0.78	1.66	3.24
		12.50		11.53	0.69	1.67	3.67
10.00	15.00	10.00	3000	3.95	0.62	1.87	5.46
		12.50		7.15	0.75	1.41	1.71
		15.00		3.65	0.65	1.61	3.65

Table IV.3.5.a The effect of approximating a discrete Poisson process with a continuous Poisson process on the decisive sampling number when $\alpha = \beta = 0.05$. This table lists $\frac{\hat{E}(N)}{E(T)}$ and $\frac{\hat{\omega}(N)}{\omega(T)}$, where $\hat{E}(N)$ and $\hat{\omega}(N)$ are estimated from simulations and $E(T)$ and $\omega(T)$ are calculated from an exact expression of the moment-generating function of the decisive sampling time T.

$\lambda_1 = (1+\Delta)\lambda_0$ Δ	λ	$\hat{E}(N)/E(T)$ $\lambda_0 =$ 1.00	5.00	10.00	$\hat{\omega}(N)/\omega(T)$ $\lambda_0 =$ 1.00	5.00	10.00
	λ_0		1.04*	1.08*		0.93*	1.04*
0.1	$(\lambda_1+\lambda_0)/2$		-	-		-	-
	λ_1		1.05*	1.10*		0.99*	1.05*
	λ_0	1.04	1.13	1.18	0.95	0.98	0.94
0.25	$(\lambda_1+\lambda_0)/2$	1.03	1.20	1.29	0.96	0.98	0.95
	λ_1	1.03	1.14	1.23	1.00	0.96	0.93
	λ_0	1.10	1.29	1.39	0.97	0.95	0.91
0.50	$(\lambda_1+\lambda_0)/2$	1.08	1.38	1.61	1.02	0.97	0.91
	λ_1	1.07	1.30	1.42	0.97	0.95	0.87

* In this case $E(T)$ and $\omega(T)$ are calculated from an approximation of the characteristic function of N.

Table IV.3.5.b The effect of approximating a discrete Poisson process with a continuous Poisson (Wiener) process on the decisive sampling number when $\alpha = \beta = 0.05$. This table lists $\frac{\hat{\gamma}_1(N)}{\gamma_1(T)}$ and $\frac{\hat{\gamma}_2(N)}{\gamma_2(N)}$, where $\hat{\gamma}_1(N)$ and $\hat{\gamma}_2(N)$ are estimated from simulations, $\gamma_1(T)$ is calculated from an exact expression of the moment-generating function of T and $\gamma_2(N)$ is calculated from an approximation to the characteristic function of N.

$\lambda_1 = (1+\Delta)\lambda_0$ Δ	λ	$\hat{\gamma}_1(N)/\hat{\gamma}_1(T)$ $\lambda_0 =$ 1.00	5.00	10.00	$\hat{\gamma}_2(N)/\hat{\gamma}_2(N)$ $\lambda_0 =$ 1.00	5.00	10.00
	λ_0		0.78*	0.90*		0.48	0.82
0.1	$(\lambda_1+\lambda_0)/2$		-	-		-	-
	λ_1		0.94*	1.00*		0.89	0.98
	λ_0		0.99	0.89		0.95	0.81
0.25	$(\lambda_1+\lambda_0)/2$		0.93	0.87		-	-
	λ_1		0.96	0.91		0.87	0.78
	λ_0		0.91	0.99		0.84	0.99
0.5	$(\lambda_1+\lambda_0)/2$		0.85	0.72		-	-
	λ_1		1.00	0.86		0.94	0.78

* In this case $\gamma_1(T)$ is calculated from an approximation of the characteristc function of N.

Table IV.4.1 Sample estimates of the average sampling number, $\hat{E}(N) = \mu_N$, with sample estimates of the standard deviation $\hat{D}(\mu_N)$ for one-sided and two-sided SPRT, and values of the average sampling number E(N) calculated from an approximate expression of the ASN-function when the risk levels are $\alpha = \beta = 0.05$.

				One-sided SPRT			Two-sided SPRT. Truncation point equal to 1.25 times the fixed-sample size		
λ_0	λ_1	λ	Number of simulations	$\hat{E}(N) = \hat{\mu}_N$	$\hat{D}(\hat{\mu}_N)$	E(N)	$\hat{E}(N) = \hat{\mu}_N$	$\hat{D}(\hat{\mu}_N)$	E(N)
5.00	5.50	5.000	1500	299.90	1.02	297.08	151.94	1.95	151.58
		5.125							191.50
		5.250		239.38	2.44		215.94	2.07	
		5.375							186.61
		5.500		125.21	2.10	112.23	154.06	2.00	140.66
5.00	6.25	5.00	3000	49.01	0.10	47.81	27.03	0.24	26.82
		5.500				40.99			34.99
		5.625		40.13	0.27		37.91	0.25	
		5.750				33.81			34.29
		6.250		21.96	0.25	18.02	26.55	0.25	22.50
5.00	7.50	5.00	3000	12.87	0.02	12.07	8.25	0.07	7.77
		6.00				10.48			9.49
		6.25		10.67	0.07		10.98	0.07	
		6.50				8.70			9.14
		7.50		6.30	0.07	4.55	7.54	0.07	5.64
10.00	11.00	10.00	1500	151.19	0.44	148.54	75.65	0.98	75.79
		10.50		120.49	1.23		109.64	0.99	
		11.00		63.12	1.05	56.12	76.82	0.99	70.33
10.00	12.50	10.00	3000	24.52	0.05	23.90	14.07	0.12	13.41
		11.00				20.49			17.50
		11.25		20.33	0.13		19.77	0.12	
		11.50				16.91			17.14
		12.50		11.42	0.12	9.06	13.73	0.12	11.25
10.00	15.00	10.00	3000	6.92	0.01	6.03	4.46	0.03	3.88
		12.00				5.24			4.75
		12.50		5.85	0.03		5.79	0.03	
		13.00				4.35			4.57
		15.00		3.53	0.04	2.27	4.08	0.03	2.82

Table IV.4.2 Ratio between sample estimates of the average sampling number, $\hat{E}(N)$, and values calculated from an approximate expression $E(N)$ when the risk levels are $\alpha = \beta = 0.05$.

$\lambda_1 = (1+\Delta)\lambda_0$ Δ	λ	$\hat{E}(N)/E(N)$ One-sided SPRT λ_0 5.00	One-sided SPRT λ_0 10.00	Two-sided SPRT truncation point equal to fixed-sample size λ_0 5.00	Two-sided SPRT truncation point equal to fixed-sample size λ_0 10.00	Two sided SPRT truncation point equal to 1.25 times the fixed-sample size λ_0 5.00	Two sided SPRT truncation point equal to 1.25 times the fixed-sample size λ_0 10.00
0.1	λ_0	1.01	1.02	0.99	0.99	1.00	1.00
	λ_1	1.12	1.12	1.03	1.03	1.10	1.09
0.25	λ_0	1.03	1.03	1.00	1.03	1.01	1.05
	λ_1	1.22	1.26	1.07	1.11	1.18	1.22
0.50	λ_0	1.07	1.15	1.03	1.01	1.06	1.15
	λ_1	1.38	1.56	1.18	1.17	1.34	1.45

Chapter V COMPARISON OF DIFFERENT TEST PROCEDURES

This paper has been aimed at studying truncated sequential probability ratio tests for different models using the assumption of a Poisson distributed stochastic variable and at comparing these tests with the SPRT (sequential probability ratio test) and the optimum fixed-sample test. The comparison is made essentially in terms of the OC (operating characteristic) and ASN (average sampling number) functions, although it is also necessary to study the variation and skewness of the decisive sampling number.

Derivation of the OC and ASN-functions for truncated tests depends on the distribution of the decisive sampling number in the SPRT, for which we only have an approximate expression. This, in turn, leads to errors in the OC and ASN-functions for truncated tests.

The error in the approximation of the distribution of the decisive sampling number in the SPRT was evaluated in Section IV.2. This was done by comparing values of the mean, variance, skewness and excess calculated from this approximation, with values calculated on the basis of expressions derived from an exact expression of the moment-generating function of the decisive sample time (Dvoretzky et al., 1953), and values estimated from simulations. It was found in Section IV.3 that the approximation underestimated the average sampling number, E(N), but gave more accurate estimates of the coefficients of variation and skewness. The underestimate of E(N) increased with λ_0 and the ratio λ_1/λ_0.

The accuracy of the approximations of the OC and ASN-functions for truncated tests is studied by comparisons with results from simulations of one-sided and two-sided truncated SPRT at intended risk levels $\alpha = \beta = 0.05$. The results from the simulations of one-sided truncated SPRT are presented in Table V.1 and those of two-sided truncated SPRT are presented in Table V.2. The ratios between sample estimates of the average sampling number and the average

sampling number calculated from the approximations are studied for different values of λ_0, λ_1 and λ. The actual risk levels and these ratios for SPRT, one-sided and two-sided truncated SPRT are presented in Table V.3. The simulations indicate that the actual risk levels for all tests are lower than or equal to the intended risk levels. The actual risk levels for the SPRT and the two-sided SPRT with the truncation point equal to 1.25 times the fixed-sample size are quite similar. The test with the truncation point equal to fixed-sample size has higher actual risk levels. The actual risk levels for the one-sided truncated SPRT are different from these of the other tests.

With regard to all tests, the ratio between the sample-estimate of the ASN-function and the approximation of the ASN-function increases with λ_0 and λ_1/λ_0. The relation between this ratio and the actual risk levels for different tests procedures is shown in Figure V.1. Only the SPRT and the two-sided truncated SPRT with the truncation point equal to 1.25 times the fixed-sample size had similar actual risk levels and can be directly compared. A comparison between them indicates a lower underestimate of the ASN-function at $\lambda = \lambda_0$ and an equally high or higher underestimate at $\lambda = \lambda_1$ in the truncated test when the approximations given in this study are used.

We now turn to a study of the differences between the different test procedures in ASN-functions and in the variance and the coefficient of skewness of the decisive sampling number, N. In order to simplify this comparison, sample estimates of these characteristics and actual risk levels are listed for the test procedures in Table V.4. The results in this table indicate that the test procedures are characterized as follows.

The SPRT is characterized by low ASN-values at λ_0 and λ_1, a positive coefficient of skewness and a high variance of N when $\lambda = (\lambda_1+\lambda_0)/2$.

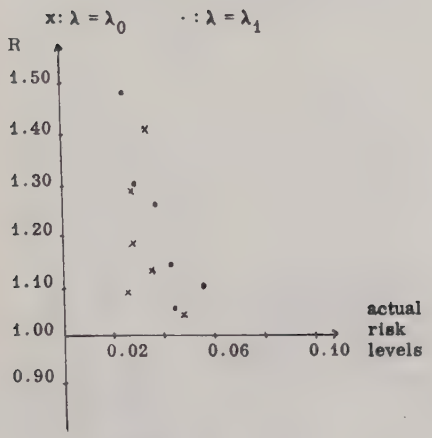
Use of a one-sided truncated SPRT can be regarded as an alternative when there is considerable confidence that λ is close to or higher than λ_1. This test then has an average sampling number that is smaller than that for the SPRT, the variance of N is about the same and will be lower if λ comes close to $(\lambda_1+\lambda_0)/2$ and the coefficient of skewness is smaller. Compared with two-sided truncated SPRT, the average sampling number is smaller for $\lambda > (\lambda_1+\lambda_0)/2$ but both the variance and the coefficient of skewness seem to be a little higher.

Truncation of the SPRT — a two-sided truncated SPRT — will not always reduce the maximum of the ASN-function but will reduce the maximal variance and the coefficient of skewness of N. A smaller truncation point leads to higher values of the ASN-function outside the region around $\lambda = (\lambda_1+\lambda_0)/2$. The variance of N will be reduced especially around $\lambda = (\lambda_1+\lambda_0)/2$ and the coefficient of skewness will decrease the most for values around $\lambda = (\lambda_1+\lambda_0)/2$. Choosing the truncation point equal to 1.25 times the fixed-sample size seems to be a good compromise, as the variance of N will be about the same for λ_0, $(\lambda_0+\lambda_1)/2$ and λ_1, and the coefficient of skewness will be best balanzed.

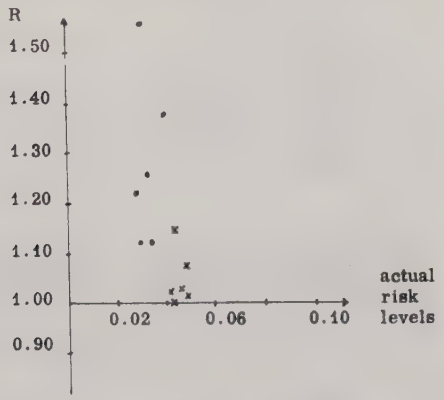
Table V.5. lists values of the ASN-functions — calculated from the approximate expression derived in this study — for two-sided truncated SPRT with $\alpha = \beta = 0.05$ and with the truncation point equal to 1.0, 1.1, 1.25 and 1.50 times the fixed-sample size.

The choice of an appropriate truncation point has to take into consideration the average sampling number and the variance and coefficient of skewness of N for those values of λ that are the most relevant in the specific problem. When chosing between tests with different truncation points, the indication that the ASN-function is underestimated less by the approximation for lower truncation points also has to be taken into account.

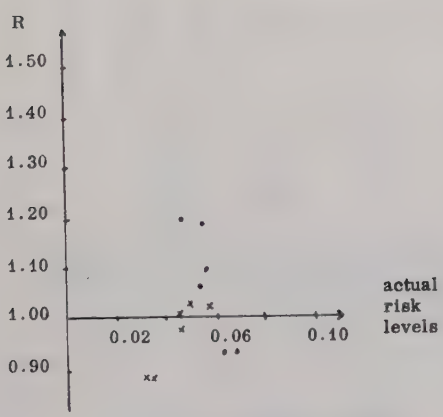
Figure V.1 Ratio, R, between $\hat{E}(N|\lambda)$ estimated from simulations and $E(N|\lambda)$ calculated from approximate expressions in relation to actual risk levels, α, β.



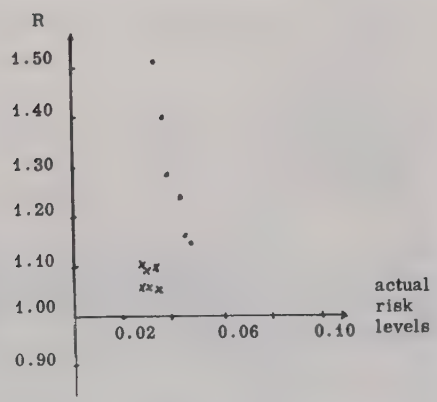
i) SPRT



ii) One-sided truncated SPRT with an upper finite limit



iii) Two-sided truncated SPRT. Truncation point equal to fixed-sample size



iv) Two-sided truncated SPRT. Truncation point equal to 1.25 times the fixed-sample size

Table V.1 Simulations of one-sided truncated SPRT

λ_0	λ_1	λ	Upper boundary to continuation region of log likelihood ratio	Truncation point N_0	Number of simulations	OC-function estimated from simulations	ASN-function: Estimated from simulations	ASN-function: Calculated from approximation
5.000	5.500	5.000	2.944	307	1500	0.959	299.90	297.08
		5.250				0.537	239.38	-
		5.500				0.032	125.21	112.23
5.000	6.250	5.000	2.944	50	3000	0.954	49.01	47.81
		5.500				-	-	40.99
		5.625				0.542	40.13	-
		5.750				-	-	33.81
		6.250				0.029	21.96	18.02
5.000	7.500	5.000	2.944	13	3000	0.962	12.87	12.11
		6.000				-	-	10.51
		6.250				0.545	10.67	-
		6.500				-	-	8.72
		7.500				0.039	6.30	4.55
10.000	11.000	10.000	2.944	154	1500	0.961	151.19	148.54
		10.500				0.545	120.49	-
		11.000				0.030	63.12	56.12
10.000	12.500	10.000	2.944	25	3000	0.958	24.54	23.90
		11.000				-	-	20.49
		11.250				0.538	20.33	-
		11.500				-	-	16.91
		12.500				0.032	11.42	9.06
10.000	15.000	10.000	2.944	7	3000	0.968	6.92	6.06
		12.000				-	-	5.25
		12.500				0.552	5.85	-
		13.000				-	-	4.35
		15.000				0.030	3.53	2.27

Table V.2 Simulations of two-sided truncated SPRT

λ_0	λ_1	λ	Number of simulations	I Truncation point equal to fixed-sample size: OC-function estimated from simulations	I: ASN-function estimated from simulations	I: ASN-function Approx.	II Truncation point equal to 1.25 times fixed-sample size: OC-function estimated from simulations	II: ASN-function estimated from simulations	II: ASN-function Approx.
5.000	5.500	5.000	1500	0.965	196.53	198.86	0.971	151.94	151.58
		5.125				215.98			191.50
		5.250		0.504	222.40		0.520	215.94	
		5.375				214.43			186.61
		5.500		0.061	197.89	191.79	0.047	154.06	140.66
5.000	6.250	5.000	3000	0.953	32.53	32.63	0.967	27.03	26.82
		5.500				36.71			34.99
		5.625		0.498	37.49		0.492	37.91	
		5.750				36.49			34.29
		6.500		0.054	31.60	29.40	0.042	26.55	22.50
5.000	7.500	5.000	3000	0.950	9.13	8.84	0.971	8.25	7.77
		6.000				9.79			9.49
		6.250		0.455	10.41		0.474	10.98	
		6.500				9.65			9.14
		7.500		0.048	8.43	7.15	0.037	7.54	5.64
10.000	11.000	10.000	1500	0.967	98.76	99.43	0.970	75.65	75.79
		10.500		0.507	111.68		0.515	109.64	
		11.000		0.067	99.17	95.90	0.044	76.82	70.33
10.000	12.500	10.000	3000	0.954	16.76	16.31	0.968	14.07	13.41
		11.000				18.36			17.50
		11.250		0.511	19.27		0.506	19.77	
		11.500				18.24			17.14
		12.500		0.055	16.34	14.70	0.039	13.73	11.25
10.000	15.000	10.000	3000	0.943	4.48	4.42	0.971	4.46	3.88
		12.000				4.90			4.75
		12.500		0.456	4.85		0.478	5.79	
		13.000				4.82			4.57
		15.000		0.054	4.20	3.58	0.032	4.08	2.82

Table V.3 Ratio, R, between $\hat{E}(N)$ estimated from simulations and $E(N)$ calculated from approximate expression and actual risk levels estimated from simulations

			SPRT		One-sided truncated SPRT		Two-sided truncated SPRT with truncation point equal to fixed-sample size		1.25 times fixed-sample size	
λ_0	λ_1	λ	Risk levels	R	Risk levels	R	Risk levels	R	Risk levels	R
5.0	5.5	5.0	0.047	1.04	0.041	1.00	0.035	0.99	0.029	1.00
		5.5	0.046	1.05	0.032	1.12	0.061	1.03	0.047	1.10
5.0	6.25	5.0	0.034	1.13	0.046	1.03	0.047	1.00	0.033	1.01
		6.25	0.041	1.14	0.029	1.22	0.054	1.07	0.042	1.18
5.0	7.5	5.0	0.027	1.29	0.048	1.07	0.050	1.03	0.029	1.06
		7.5	0.028	1.30	0.039	1.38	0.048	1.18	0.037	1.34
10.0	11.0	10.0	0.025	1.08	0.049	1.02	0.033	0.99	0.030	1.00
		11.0	0.054	1.10	0.030	1.12	0.067	1.03	0.044	1.09
10.0	12.5	10.0	0.027	1.19	0.042	1.03	0.046	1.03	0.032	1.05
		12.5	0.037	1.26	0.032	1.26	0.055	1.11	0.039	1.22
10.0	15.0	10.0	0.033	1.41	0.042	1.15	0.057	1.01	0.029	1.15
		15.0	0.023	1.49	0.030	1.56	0.054	1.17	0.032	1.45

Table V.4 Values of actual risk levels, average sampling number and standard deviation and coefficient of skewness of decisive sampling number (N) estimated from simulations of SPRT, one-sided truncated SPRT and two-sided truncated SPRT

λ_0	λ_1	λ	Estimated risk levels				Average sampling number $E(N) = \mu_N$				Standard deviation of N				Coefficient of skewness of N			
				Truncated				Truncated				Truncated				Truncated		
				One-sided	Two-sided SPRT*			One-sided	Two-sided SPRT*			One-sided	Two-sided SPRT*			One-sided	Two-sided SPRT*	
			SPRT	SPRT	I	II	SPRT	SPRT	I	II	SPRT	SPRT	I	II	SPRT	SPRT	I	II
5.000	5.500	5.000	0.047	0.041	0.035	0.029	117.01	299.90	196.53	151.94	77.9	39.6	42.4	75.6	1.5	-5.6	-1.2	0.5
		5.250					198.42	239.38	222.40	215.94	161.3	94.6	18.9	80.0	1.6	-0.9	-4.8	-0.7
		5.500	0.049	0.032	0.061	0.047	114.41	125.21	197.89	154.06	83.1	81.5	41.8	77.4	1.8	1.0	-1.3	0.4
5.000	6.250	5.000	0.038	0.046	0.047	0.033	22.40	49.01	32.53	27.03	15.3	5.6	7.9	12.9	1.9	-6.0	-0.9	0.4
		5.625					38.10	40.13	37.49	37.91	30.6	14.7	4.7	13.5	1.8	-1.1	-3.4	-0.8
		6.250	0.040	0.029	0.054	0.042	21.57	21.96	31.60	26.55	15.3	13.8	8.9	13.6	1.8	0.7	-0.9	0.4
5.000	7.500	5.000	0.030	0.038	0.050	0.029	7.27	12.87	9.13	8.25	4.7	1.1	2.3	3.6	1.8	-8.5	-0.9	0.3
		6.250					12.33	10.67	10.41	10.98	9.7	3.6	1.6	3.7	1.7	-1.2	-2.9	-0.8
		7.500	0.033	0.039	0.048	0.037	6.70	6.30	8.43	7.54	4.8	3.7	2.7	3.9	1.6	0.6	-0.6	0.4
10.000	11.000	10.000	0.025	0.039	0.033	0.030	60.93	151.19	98.76	75.65	41.9	17.2	21.4	37.9	1.7	-6.4	-1.2	0.5
		10.500					102.87	120.49	111.68	109.64	82.5	47.7	9.4	38.4	1.7	-1.0	-4.7	-0.8
		11.000	0.054	0.030	0.067	0.044	60.34	63.12	99.17	76.82	42.0	40.8	21.3	38.5	1.9	0.9	-1.2	0.4
10.000	12.500	10.000	0.027	0.042	0.046	0.032	11.71	24.52	16.76	14.07	7.7	2.8	4.0	6.5	1.7	-5.9	-0.9	0.4
		11.250					20.60	20.33	19.27	19.77	16.0	7.2	2.3	6.7	1.7	-1.2	-3.6	-0.9
		12.500	0.037	0.032	0.055	0.039	11.53	11.42	16.32	13.73	8.0	6.8	4.5	6.8	1.7	0.7	-0.9	0.4
10.000	15.000	10.000	0.033	0.032	0.057	0.029	3.95	6.92	4.48	4.46	2.43	0.61	0.85	1.77	1.9	-8.2	-1.5	0.1
		12.500					7.15	5.85	4.85	5.79	5.34	1.86	0.52	1.71	1.4	-1.3	-3.9	-1.1
		15.000	0.023	0.030	0.054	0.032	3.65	3.53	4.20	4.08	2.39	1.96	1.06	1.91	1.6	1.6	-1.1	0.2

* Two-sided SPRT I: Truncation point equals fixed-sample size
II: Truncation point equals 1.25 times fixed-sample size

Table V.5 Values of boundaries to the continuation region and the ASN-function for two-sided truncated SPRT for different truncation points when the intended risk levels are $\alpha = \beta = 0.05$.

			Truncation point equal to											
			fixed-sample size			1.10 times fixed-sample size			1.25 times fixed-sample size			1.50 times fixed-sample size		
λ_0	λ_1	λ	a	N_0	E(N)	a	N_0	E(N)	a	N_0	E(N)	a	N_0	E(N)
1.00	1.10	1.000	6.519	1136	994	4.600	1249	840	3.903	1420	757	3.448	1703	689
		1.025			1080			1007			957			905
		1.075			1072			989			933			875
		1.100			959			789			703			633
		1.150			617			439			373			329
		1.200			399			281			239			211
1.00	1.25	1.00	5.717	194	167	4.460	213	146	3.846	242	134	3.423	291	124
		1.05			176			166			159			150
		1.10			184			179			175			169
		1.15			182			177			171			164
		1.20			170			157			146			136
		1.25			147			126			112			101
		1.40			74			58			50			45
		1.50			51			40			34			31
1.00	1.50	1.00	5.183	53.1	44.2	4.303	58.4	41.2	3.776	66.4	38.8	3.390	79.6	36.5
		1.10			47.1			45.5			44.1			42.4
		1.20			49.0			48.3			47.5			46.2
		1.30			48.2			47.0			45.7			43.9
		1.40			43.7			40.7			38.1			35.3
		1.50			35.8			31.4			28.2			25.4
		1.75			18.0			15.0			13.1			11.8
		2.00			10.5			8.7			7.7			6.9

Chapter VI EFFECT OF HETEROGENEITY IN PARAMETER VALUES

It has been assumed in this study that there are logical grounds for the assumption that the random variables follow a Poisson distribution. Several ways of allowing for the fact that the intensity can vary between observations were discussed in Section I.6. Then two models — the negative binomial distribution and the multiplicate Poisson model — were studied in Chapter III. It was also indicated in Section I.6 and in the examples in Chapter VII how a reasonable model could be found. We begin by trying to divide the data into more homogeneous groups on the basis of some criteria variables. This can be done using the AID-method (Sonquist, Baker and Morgan (1971)) or by means of some other method. As the aim is to obtain homogeneous groups, it then seems reasonable to find groups which are constructed in such a way that the between-group variance will be maximized given the number of groups. This method implies that the variance within the groups will be underestimated. When deciding on the number of groups, several factors have to be taken into account, such as the number of observations in a group, determining which tests can be used with regard to the composition of the groups, i.e. the effects of assigning observations to a particular (see Section VII.3 for an example), and the influence possible heterogeneity in a group will have on tests.

One way of introducing a variation in λ is to regard it as a stochastic variable and then construct a model with some compound distribution. The simplest one is the negative binomial distribution, which has been used in connection with traffic accidents; see e.g. Greenwood & Yule (1920), Newbold (1927) and Arbous & Kerrich (1951). SPR-tests of the parameter μ when δ is known in a negative binomial distribution with mean δ/μ and variance $\frac{\delta}{\mu}(1+\frac{1}{\mu})$ are studied in Section III.4. In order to fit the observations to a negative binomial distribution, the variation in the intensity should not be related to the order in which the observations will be studied in the sequential test. It then becomes important to study whether known facts indicate such dependence and to take this dependence into accout.

The effects of a variation in λ can be examined by studying the negative binomial distribution. The stochastic variable X follows this distribution if X given λ is Poisson distributed and if λ follows a gamma distribution, making a transformation so that the mean δ/μ equals the constant value $\bar{\lambda}$, i.e. the mean of λ, and the variance of λ equals $\bar{\lambda}^2/\delta$. The larger the value of δ, the smaller the variance in the gamma distribution and as a limit when $\delta \to \infty$ we get the Poisson distribution. It is now interesting to see what happens to the OC and ASN-functions for the fixed-sample size test, the SPRT and the two-sided truncated SPRT. OC-functions of a SPRT for selected values of λ_0, λ_1, α and β are given in Table VI.1. Approximate risk levels for fixed-sample size tests for a Poisson parameter when the true distribution is negative binomial are shown in Table VI.2. These tables indicate that incorrect use of the Poisson distribution does not disturb the OC-function to any great extent for the sequential test even if the variance of λ is large, while the risk levels for the fixed-sample size test will be considerably higher even if the variance of λ is moderate. (See e.g. the case $\alpha = \beta$, $\lambda_0 = 10.0$ and $\lambda_1 = 11.0$ or 15.0.) In order to compare the efficiency of the test, we can study the ASN-function computed with intended risk levels α and β for SPRT (Table VI.3) and two-sided truncated SPRT (Table VI.4) for two truncation points, in one case equal to the fixed-sample size and in the other equal to 1.25 times the fixed-sample size. Table VI.5 lists the fixed-sample sizes necessary to keep the risk levels equal to α and β. The maximal rise in the ASN-function when $\lambda_0 \leq \lambda \leq \lambda_1$ as a per cent of the ASN-function for the test of the Poisson parameter is summarized in Table VI.6 when $\delta = 5$, 10 and 20. Values of the ASN-function when $\lambda_0 \neq 1.0$ for a given ratio λ_1/λ_0 are given by

$$\frac{E(N;\lambda|\lambda_0=1.0)}{\lambda_0}$$

observing that δ now equals λ_0 times the value of δ, when $\lambda_0 = 1.0$.

Table VI.1 Values of the OC-function in a SPRT of a negative binomial distribution with different values of δ compared with the OC-function in a SPRT of a Poisson distribution

i) $\lambda_0 = 1.0$ $\lambda_1 = 1.1$ $\alpha = \beta = 0.05$

δ \ λ	λ_0				λ_1
	1.000	1.025	1.050	1.075	1.100
1.0	0.9564	0.8102	0.4823	0.1844	0.0510
2.0	0.9557	0.8104	0.4840	0.1844	0.0561
3.0	0.9954	0.8106	0.4850	0.1845	0.0557
5.0	0.9550	0.8108	0.4860	0.1846	0.0533
10.0	0.9546	0.8106	0.4863	0.1844	0.0548
Poisson case	0.9544	0.8114	0.4883	0.1851	0.0545

ii) $\lambda_0 = 1.0$ $\lambda_1 = 1.5$ $\alpha = \beta = 0.05$

δ \ λ	λ_0				λ_1
	1.000	1.125	1.250	1.375	1.500
1.0	0.9738	0.7919	0.4256	0.1833	0.0849
2.0	0.9719	0.7952	0.4333	0.1829	0.0802
3.0	0.9708	0.7968	0.4375	0.1827	0.0778
5.0	0.9697	0.7984	0.4418	0.1826	0.0754
10.0	0.9686	0.7998	0.4460	0.1826	0.0730
Poisson case	0.9673	0.8016	0.4520	0.1827	0.0702

iii) $\lambda_0 = 10.0$ $\lambda_1 = 11.0$ $\alpha = \beta = 0.05$

δ \ λ	λ_0				λ_1
	10.00	10.25	10.50	10.75	11.00
1.0	0.9580	0.8092	0.4777	0.1839	0.0540
2.0	0.9577	0.8093	0.4784	0.1840	0.0586
3.0	0.9575	0.8095	0.4791	0.1840	0.0583
5.0	0.9571	0.8098	0.4804	0.1842	0.0578
10.0	0.9564	0.8101	0.4821	0.1843	0.0570
Poisson case	0.9444	0.8115	0.4884	0.1851	0.0545

Table VI.1 /cont./

iv) $\lambda_0 = 1.0 \quad \lambda_1 = 1.1 \quad \alpha = \beta = 0.01$

δ \ λ	λ_0 1.000	1.025	1.050	1.075	λ_1 1.100
1.0	0.9920	0.9059	0.4725	0.0894	0.0124
2.0	0.9918	0.9061	0.4751	0.0895	0.0121
3.0	0.9917	0.9063	0.4766	0.0896	0.0119
5.0	0.9916	0.9065	0.4782	0.0897	0.0118
10.0	0.9914	0.9063	0.4787	0.0895	0.0114
Poisson case	0.9914	0.9070	0.4818	0.0900	0.0115

Table VI.2 Approximate risk levels with fixed-sample size test if the test for a Poisson parameter is used when the correct distribution is negative binomial

	$\lambda_0 = 1.0$				$\lambda_0 = 10.0$			
	$\lambda_1 = 1.1$		$\lambda_1 = 1.5$		$\lambda_1 = 11.0$		$\lambda_1 = 15.0$	
δ	α	β	α	β	α	β	α	β
1.0	0.12	0.13	0.12	0.15	0.31	0.32	0.31	0.34
2.0	0.09	0.11	0.09	0.10	0.25	0.26	0.25	0.29
3.0	0.08	0.09	0.08	0.09	0.21	0.22	0.21	0.25
5.0	0.07	0.07	0.07	0.07	0.17	0.18	0.17	0.20
10.0	0.06	0.06	0.06	0.06	0.12	0.13	0.12	0.15
Poisson	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05

Table VI.3 Values of the ASN-function in a SPRT of a negative binomial distribution with different values of δ compared with the ASN-function in a SPRT of a Poisson distribution

i) $\lambda_0 = 1.0$ $\lambda_1 = 1.1$ $\alpha = \beta = 0.05$

δ \ λ	λ_0 1.000	1.025	1.050	1.075	λ_1 1.100
1.0	1184	1650	1860	1525	1095
2.0	878	1223	1385	1139	818
3.0	775	1080	1226	1011	726
5.0	693	966	1099	908	652
10.0	632	881	1004	831	597
Poisson case	571	795	909	754	542

ii) $\lambda_0 = 1.0$ $\lambda_1 = 1.5$ $\alpha = \beta = 0.05$

δ \ λ	λ_0 1.000	1.125	1.250	1.375	λ_1 1.500
1.0	68.35	95.32	92.05	67.71	48.57
2.0	48.76	67.98	67.18	50.13	36.00
3.0	42.22	58.84	58.85	44.26	31.80
5.0	36.98	51.52	52.16	39.56	28.44
10.0	33.05	46.02	47.12	36.03	25.92
Poisson case	29.11	40.50	42.05	32.49	23.39

iii) $\lambda_0 = 1.0$ $\lambda_1 = 1.1$ $\alpha = \beta = 0.01$

δ \ λ	λ_0 1.000	1.025	1.050	1.075	λ_1 1.100
1.0	1992	3370	4528	3096	1881
2.0	1478	2497	3371	2313	1404
3.0	1306	2205	2985	2053	1245
5.0	1169	1972	2676	1845	1118
10.0	1067	1798	2444	1687	1022
Poisson case	963	1621	2213	1533	927

Table VI.4 Values of the ASN-function in a truncated SPRT of a negative binomial distribution with different values of δ compared with the ASN-function in a truncated SPRT of a Poisson distribution

i) $\lambda_0 = 1.00 \quad \lambda_1 = 1.10 \quad \alpha = \beta = 0.05$
$N_0 = 1136.16$ = fixed-sample size

δ \ λ	λ_0				λ_1		
	0.950	1.000	1.025	1.075	1.100	1.150	1.200
1.0	1135	1136	1136	1136	1136	1135	1130
2.0	1087	1129	1134	1134	1128	1082	955
3.0	1019	1115	1130	1129	1111	993	783
5.0	928	1088	1120	1117	1077	871	622
10.0	837	1052	1106	1101	1031	750	505
20.0	785	1026	1095	1089	999	684	450
Poisson case	731	994	1080	1072	959	617	399

ii) $\lambda_0 = 1.00 \quad \lambda_1 = 1.10 \quad \alpha = \beta = 0.05$
$N_0 = 1420.20 = 1.25$ times fixed-sample size

δ \ λ	λ_0				λ_1		
	0.950	1.000	1.025	1.075	1.100	1.150	1.200
1.0	1233	1366	1393	1392	1367	1260	1092
2.0	895	1196	1293	1285	1183	884	632
3.0	745	1084	1218	1206	1060	709	481
5.0	624	970	1134	1118	934	567	374
10.0	535	869	1054	1034	824	466	302
20.0	429	814	1007	985	764	418	270
Poisson case	451	757	957	933	703	373	239

iii) $\lambda_0 = 1.00 \quad \lambda_1 = 1.50 \quad \alpha = \beta = 0.05$
$N_0 = 53.08$ = fixed-sample size

δ \ λ	λ_0				λ_1		
	0.750	1.000	1.125	1.375	1.500	1.750	2.000
1.0	52.5	53.1	53.1	53.1	53.1	53.0	52.8
2.0	48.8	52.3	52.8	52.6	51.9	48.4	42.5
3.0	46.0	51.1	52.2	51.7	49.7	41.4	31.5
5.0	43.1	49.3	51.2	50.2	46.1	32.6	21.9
10.0	40.8	47.2	49.9	48.2	41.7	25.1	15.7
20.0	39.6	45.8	48.9	46.8	39.0	21.5	12.9
Poisson case	38.4	44.2	47.7	45.2	35.8	18.0	10.5

Table VI.4 /cont./

iv) $\lambda_0 = 1.00$ $\lambda_1 = 1.50$ $\alpha = \beta = 0.05$
$N_0 = 66.35 = 1.25$ times fixed-sample size

δ \ λ	λ_0			λ_1			
	0.750	1.000	1.125	1.375	1.500	1.750	2.000
1.0	57.0	64.6	65.6	65.4	64.8	62.7	59.8
2.0	44.5	58.0	61.6	60.4	56.2	44.5	34.1
3.0	39.4	53.3	58.2	56.1	49.3	33.7	23.3
5.0	35.5	48.4	54.3	51.1	41.9	24.7	16.0
10.0	32.6	43.9	50.3	46.3	35.3	18.5	11.4
20.0	31.3	41.4	47.9	43.4	31.8	15.7	9.4
Poisson case	30.1	38.8	45.2	40.4	28.2	13.1	7.7

Table VI.5 Approximate fixed-sample size when X is Poisson or negative binomially distributed and the risk levels are $\alpha = \beta = 0.05$

δ	$\lambda_0 = 1.0$		$\lambda_0 = 10.0$	
	$\lambda_1 = 1.1$	$\lambda_1 = 1.5$	$\lambda_1 = 11.0$	$\lambda_1 = 15.0$
1.0	2203	120	1299	73
2.0	1722	87	706	38*
3.0	1520	76	508	28*
5.0	1362	69	350	19*
10.0	1240	60	229	12*
Poisson	1136	53	114	6

* Approximate values

Table VI.6 Approximate maximal percentual rise in the ASN-function when X follows a negative binomial distribution with $E(\lambda) = \bar{\lambda}$ and $V(\lambda) = \bar{\lambda}^2/\delta$ as a per cent of the ASN-function when X follows a Poisson distribution with the parameter $\bar{\lambda}$ in the interval $\lambda_0 \leq \lambda \leq \lambda_1$

λ_0	λ_1	δ	Fixed-sample size test	SPRT	Two-sided truncated SPRT truncation point equal to 1.25 times the fixed-sample size	Two-sided truncated SPRT truncation point equal to fixed-sample size
1.0	1.1	5	20	20	35	13
		10	10	11	20	8
		20			10	5
1.0	1.5	5	30	30	50	30
		10	15	15	25	17
		20			13	9

Chapter VII ON THE APPLICATION OF SEQUENTIAL TESTS IN THE ANALYSIS OF ROAD TRAFFIC ACCIDENTS

1. Introduction

The two examples presented in this chapter show how sequential analysis could be applied to data from the changeover to right-hand traffic in Sweden September 3, 1967. The actual analysis was performed without using sequential test procedures, although some of the decision procedures had sequential aspects.

At the time of the changeover, the Right-hand Traffic Commission's policy was that after the changeover, traffic safety should not decrease for any traffic category. It was then necessary to construct some measure of the traffic safety level. Data about traffic accidents of varying severity and for different traffic categories in the whole country were used and related to accidents during a reference period. Congested and non-congested areas were seperated in the analysis. The examples below refer to congested areas. In these areas only the variation in the number of accidents according to weekday and season could be taken into account, as no measure of the amount of traffic was available.

A serious drawback was that only the year 1966 could be used as a reference year because a revision of the statistics on road traffic accidents was effectuated as of January 1, 1966. From the beginning of 1967, early changes in traffic rules and a greater amount of information on traffic safety could have changed the safety level. The decision was made to use only data from 1966 for reference, but to use other data to test the assumptions in the models that were accepted. Later research (Jörnrup & Svensson, 1971) indicates that a change in the number of road accidents reported to the police occurred in connection with the revision of the statistics. They studied the accidents in non-congested areas during 1961-1966 by means of a multiplicative Poisson model. This model indicated a decline in the number of accidents reported from 1965 to 1966. It seems obvious that this decrease is to some extent due to the revised reporting procedure. There was no similar analysis made for congested areas, but the

results for non-congested areas indicate that, for certainty, only 1966 should be used as a reference period.

Decisions related to the changeover had to be based not only on the total number of accidents but also on the distribution of these accidents into different categories. It was decided that the severity of the accident and which traffic categories were involved in the accident should be taken into consideration. The final decision model contained the following variables, with the calendar day as the primary time unit:

i) accidents with personal injury
ii) fatal accidents
iii) accidents with serious personal injury
iv) accidents of the type motor vehicle — motor vehicle, with personal injury
v) accidents of the type motor vehicle — bicycle/moped, with personal injury
vi) accidents of the type motor vehicle — pedestrian, with personal injury.

When using these variables as bases for decisions, the variables' relevance and mutual dependence have to be taken into account. With respect to relevance it is known (Rosmark & Fräki 1966) that a large number of accidents are not reported. Nearly all fatal accidents are assumed to be reported, but hospital statistics show that even many severe accidents are not reported. Hospital statistics also indicate that the proportion of reported accidents varies for different traffic categories. As the comparisons use the number of reported accidents during a reference period, this problem can be overcome as long as the proportion of reported accidents for each group remains unchanged. Still, this assumption should be kept in mind when drawing conclusions. Two of the variables above are studied separately below.

2. Fatal Accidents in Congested Areas

When looking at the number of fatal accidents in Sweden during a year, the number of fatal accidents per calendar day varies with the amount of traffic and weather conditions. Owing to the high costs involved, it was not possible to

obtain any estimate of the amount of traffic in congested areas. As the amount of traffic is assumed to differ with the weekday and season of the year and weather conditions differ with the season; weekday and month were used as dependent variables. These two variables provide a partial description of the regular variation in the amount of traffic and weather conditions during the year. These variables are used to construct time periods which are homogeneous in the sense that the number of fatal accidents during a given period approximately follows a Poisson distribution. Given the number of periods, it is then reasonable to choose them with reference to the dependent variables in such a way as to minimize the weighted variation within period. Given this criterion, different techniques could be used. During the changeover to right-hand traffic a technique was used that seems intuitively plausible. Morgan & Sonquist (1964) formalized it as AID-analysis. Later on it was developed further on by Sonquist, Baker & Morgan (1971). Gavatin-Avén (1973) have shown that the problems of significance encountered when using the AID technique are complicated. The AID technique can be described simply as follows.

With the month (weekday) as the dependent variable, two periods are constructed, which give the highest possible variance between periods without any restriction that the period must involve adjacent months (weekdays). Then, from these two classifications, the one giving the highest variance between periods is chosen. The same procedure is used in each of the two chosen periods and so on.

When deciding how far the devision is to be carried out, it should be kept in mind that each period must contain a reasonable number of days. If the number of observations are too small, a low level precision will be obtained in the estimate of the parameter used to measure the fit of the observations to a Poisson distribution. The period must also be long enough so that a truncated SPRT can be completed during the period.

As we have not used any restriction that adjacent weekdays and months have to be grouped together the logical consistency of the classification can be studied through the following facts:

a) how the weekdays and months belonging to a period in the final classification are adjacent

b) in which way the weekdays and months, where a priori knowledge about traffic characteristics and weather conditions indicate that they should be grouped together, belong to the same period in the final classification.

When we applied AID analysis to the number of fatal accidents per calendar day for the year 1966, we obtained the following two periods:

i) November

ii) January-October and December.

As this classification only explained 0.03 of the total variance and there were no reasons for seperating the month of November, the whole year was used as a single period. The proportion explained, 0.03, seemed to be very low and this should be kept in mind throughout the examination of the data.

In order to study whether it can be assumed that the observations follow a Poisson distribution with the same intensity for all observations, as opposed to the alternative that the intensities differ between observations, we study the value of

$$\hat{cv}(\lambda) = \frac{\hat{V}(x) - \bar{x}}{\bar{x}} \qquad \text{(VII.2.1)}$$

We find that $\hat{cv}(\lambda)$ equals 0.02. The heterogeneity of the intensities was studied on the basis of a negative binomial distribution and the effect of heterogeneity was studied through the parameter δ (cf. Chapter VI). If the observations are fitted to a negative binomial distribution and the parameter δ is estimated, our estimate of δ will equal 61, which means it is reasonable that a Poisson distribution could be used without difficulty. As shown in Chapter VI, if the variation in the parameter values is as small as this, then the change in both the OC-function and the ASN-function will not be large.

We have the possibility of matching observations which take into account the weekday, date and holidays. But as we have not found any clear connection between the number of fatal accidents per day and the weekday and month, and as aggregated data for the entire country with respect to the amount of traffic and weather conditions seem to explain only a small part of the variation, the advantages will not be large enough to compensate for the more complex calculations required to perform these tests.

Instead we use the before and efter experiment situation with 365 observations during the before-period. A reasonable way to choose the minimum value of the H_1-hypotheses and maximum value of the H_0-hypotheses is to construct a confidence interval around the estimate of λ based on the data of 1966 and to use the endpoints of such an interval. Using a confidence level of 0.99, we obtain the confidence interval (0.90-1.18) and, as hypotheses, we could take

$H_0: \lambda \leq 0.9$ against $H_1: \lambda \geq 1.2$.

However, before we can decide on the hypothesis, the truncation point and risk levels have to be determined. Since it was assumed that the safety level might not remain the same during long periods after the changeover, the test had to have a low truncation point so that decisions with certainty could be reached quickly. Table VII.1 shows the truncation point for one-sided and two-sided truncated SPRT when the truncation point equals 1.25 times the fixed-sample size and Table VII.2 contains the average sampling number when $\lambda = \lambda_1$. The values for the truncation points and the average sampling number are taken from Appendix F. The concentrations on the value of the ASN-function at $\lambda = \lambda_1$ depend on the fact that a high accident frequency should be found as quickly as possible.

The one-sided truncated SPRT gives both a lower truncation point and a lower average sampling number when $\lambda = \lambda_1$. If the risk levels are chosen as $\alpha = \beta = 0.05$ and the ratio $\lambda_1/\lambda_0 = 1.50$, both tests will certainly terminate

after about two months. If the ratio λ_1/λ_0 is reduced to 1.25, this period will increase to about 8 months. This period is too long to work with. Lower risk levels will not increase the truncation point as much, but from the users' point of view $\alpha = \beta = 0.05$ was the best choice.

When this choice is applied, there is little difference between one-sided and two-sided SPRT with regard to the truncation point and the average sampling number when $\lambda = \lambda_1$. As the intention was to look at the hypotheses symmetrically, the two-sided test seemed most appropriate. In order to choose parameter values of the hypotheses, a confidence interval for the average number of accidents per day during 1966 is constructed in such a way that the upper limit of the interval will equal 1.5 times the lower limit. The lower limit will then be used as the upper limit of the H_0-hypotheses and the upper limit as the lower limit of the H_1-hypotheses. This give us the hypotheses $H_0: \lambda \leq 0.83$ and $H_1: \lambda \geq 1.25$. The risk levels $\alpha = \beta = 0.05$ are given under the assumption that — after the truncation point is reached — we decide on H_1 if $S_{N_0} > 0$, otherwise on H_0. At the time of the changeover, it was more realistic to use another decision rule, i.e. decide on H_1 if $S_{N_0} \geq a$ (the traffic safety level has increased) decide on neither H_1 nor H_0 if $b < S_{N_0} < a$ (the traffic safety level is unchanged) decide on H_0 if $S_{N_0} \leq b$ (the traffic safety level has decreased). Use of this decision rule will decrease the risk levels. The extent to which this occurs can be determined from Gosh (1970). He gives lower bounds to the probability $P(N < N_0)$ (pp. 142-43) and an approximation of the probability $P(0 < S_{N_0} < a)$ (p. 222). Calculations then show that α will decrease by less than 0.009 and β by less than 0.017. In the actual analysis of road traffic accidents in congested areas at the time of the changeover to right-hand traffic, tests of fixed-sample sizes were used in the decision rules. The calendar day, the week and a four-week period with fixed starting-point were used as sampling periods. A new sampling period began as soon as the preceding period ended. If z_i, y_j and z_k, respectively, denote the number of fatal accidents during day i,

week j and four-week period k, the following decision rules were used which gave probabilities close to 0.05 for each test performed in order to decide on a decrease (increase) in the safety level for the reference period.

In order to find decision rules for the day-by-day analyses, values of the 5 per cent, P_5, and the 95 per cent, P_{95} percentiles were estimated from the reference data from 1966. The following were then used as the decision rule:

if $x_i < P_5$ decide on an "increased safety level"

if $P_5 \leq x_i \leq P_{95}$ decide on "no change"

if $x_i > P_{95}$ decide on a "decreased safety level".

There were 130 days with zero fatal accidents during 1966, so that the decision "increased safety level" could never be reached on the basis of the day-by-day analysis.

In the week-by-week analysis the total number of fatal accidents (y_j) was tested against the expected number of fatal accidents during a week estimated from 1966, that is $\lambda = 7.29$ and we obtained the decision rule:

if $y_j \leq 1$ decide on an "increased safety level"

if $1 < y_j < 15$ decide on "no change"

if $y_j \geq 15$ decide on a "decreased safety level".

For four-week-periods we used the same procedure to test the total number of accidents during the period (z_k) against the estimated expected number of accidents, 29.1. Using the normal approximation, we obtained the decision rule:

if $z_k \leq 20$ decide on an "increased safety level"

if $20 < z_k < 39$ decide on "no change"

if $z_k \geq 39$ decide on a "decreased safety level".

These decision rules were also compared with a decision rule based on a fixed-sample size test, which contained 76 observations. The results of the various tests on data about the number of fatal accidents in congested areas during the

Autumn of 1967 (Sept. 4 - Dec. 24) were studied. For fatal accidents during the Autumn of 1967, the day-by-day analysis always gave the decision "no change" and the week-by-week analysis gave the decision an "increased safety level" during the first week, otherwise the decision "no change".

Table VII.3 lists the datas when the truncated SPRT, the fixed-sample size test and the four-week tests terminate and the resulting decisions.

As these results only serve as an example, it is important not to draw any general conclusion about how the different tests work. The truncated SPRT and the fixed-sample size test are both based on the same statistical formalization of the problem. In Chapter V it was also found that the actual risk levels in the truncated SPRT were lower than the intended ones. The statistical formalization of the problem for the four-week test was carried out as follows. Test whether the number of fatal accidents differs from the expected number of fatal accidents given by the estimated number of fatal accidents during 1966 using a symmetrical two-sided test on a risk level equal to 0.10.

During the period Sept. 4 - Dec. 24, 1967, the average number of fatal accidents per day was 0.65 and for the four four-week periods 0.50, 0.64, 0.57 and 0.89, respectively.

3. Accidents with Personal Injury between Vehicles and Bicycles or Mopeds in Congested Areas

An AID analysis with the month and weekday as dependent variables was performed in order to find homogeneous or nearly homogeneous groups for testing the hypotheses. The partition was terminated when it was found that further partition would explain less than 5 per cent of the total variance. The groups obtained in this way are presented in Table VII.4. This table also contains the values of

$$\frac{\hat{V}(x) - \bar{x}}{\bar{x}} \qquad \text{(VII.3.1)}$$

for each group and the estimated values of δ which were obtained by fitting the observations in each group to a negative binomial distribution. The estimate $\hat{\delta}$ is given by

$$\delta = \frac{\bar{x}^2}{\hat{V}(x) - \bar{x}} \qquad \text{(VII.3.2)}$$

The results in Chapter VI were then used to decide whether a test could be carried out based on the assumption that the number of accidents during a period follows a Poisson distribution with the same intensity for all observations. The different groups are also presented in Figure VII.1. Some of the main findings were as follows:

i) The value of $\hat{\delta}$ for the period January - February is too low to allow for a fit to a Poisson distribution (see Chapter VI). Further partition — also presented in Table VII.4 still does not give any groups that are homogeneous enough.

ii) The variance of the observations for Mondays through Fridays during August - October is lower than the mean, giving $(\hat{V}(x) - \bar{x}) = 0.02 \cdot \bar{x}$ (a low value).
The observations fit to a Poisson distribution are tested using the dispersion test described in Section I.6 and given by the test-statistic

$$D = \frac{\Sigma (x_i - \bar{x})^2}{\bar{x}} \quad , \quad \text{where} \quad \bar{x} = \frac{\Sigma x_i}{n} \quad . \qquad \text{(VII.3.3)}$$

According to Fisher (1950), the distribution of this statistic under the hypothesis that all observations x_i are Poisson distributed with the same intensity can be approximated by a χ^2-distribution with $(n-1)$ degrees of freedom. In this case, we get a χ^2-value of 63.54 corresponding to a P-value between 0.40 and 0.50. As the value of

$$\frac{V(\bar{x}) - \bar{x}}{\bar{x}}$$

were very close to zero and as the hypotheses of the same intensity for the observations were not rejected, a Poisson distribution with homogeneous intensities are used.

iii) The value of $\hat{\delta}$ for Saturdays and Sundays during March, April and December is too low to assume that the intensity variations are small.

iv) Saturdays during May - November gave a value of $\hat{\delta}$ that is much too low in order to use a Poisson distribution with the same intensity for all

observations as an approximation. Further partition gives one period that could be more homogeneous, but the periods are too small to be separated.

When chosing a combination of test procedures the following aspects are adopted:

- o weekdays should be separated from Saturdays and Sundays, since the changeover could lead to different relative changes in the accident frequencies for these two periods
- o tests that condition a homogeneous group should be chosen only if the period contains so many observations that a truncated SPRT can be terminated
- o chose $\alpha = \beta = 0.05$
- o the ratio λ_1/λ_0 should be chosen as small as possible.

If these aspects are applied, a two-sided truncated SPRT that conditions homogeneous groups with $\lambda_1/\lambda_0 = 1.25$ could be used for the periods

i)	March, April	Monday - Friday
ii)	May - July	Monday - Friday
iii)	August - October	Monday - Friday.

In addition, a two-sided SPRT that conditions a multiplicative Poisson model is used for the two periods

i)	January - February, November - December	Monday - Friday
ii)	All year	Saturday - Sunday.

In the latter case, a truncated SPRT is not used because truncation points were not calculated for this situation in this study.

The value of λ_0 is chosen in such a way that $(\lambda_1+\lambda_0)/2$ equals the observed λ-value. This way of choosing λ_0 assumes symmetry in the loss function. However, this form of symmetry was chosen at the time of the changeover.

The following hypotheses are then used in the tests:

March - April, Monday - Friday	$H_0: \lambda \leq 4.9$	$H_1: \lambda \geq 6.1$
May - July, Monday - Friday	$H_0: \lambda \leq 7.8$	$H_1: \lambda \geq 9.8$
August - October, Monday - Friday	$H_0: \lambda \leq 9.7$	$H_1: \lambda \geq 12.1$
Other periods	$H_0: p \leq 0.44$	$H_1: p \geq 0.56$

In order to decide the truncation point for truncated SPRT assuming a Poisson distribution, we make use of the fact that the truncation point has been chosen as a constant times the fixed-sample size. The fixed-sample size equals

$$\frac{\chi^2_{2(f+1),\alpha}}{2\lambda_0} \quad , \qquad \text{(VII.3.4)}$$

where $f = f(\alpha, \beta)$ is the smallest integer such that

$$\frac{\chi^2_{2(f+1),\,1-\beta}}{\chi^2_{2(f+1),\,\alpha}} \leq \frac{\lambda_1}{\lambda_0} \quad . \qquad \text{(VII.3.5)}$$

For given risk levels and a given ratio λ_1/λ_0 the average sampling number is given by the function

$$E(N;\lambda) = \frac{k}{\lambda_0} \qquad \text{(VII.3.6)}$$

where k equals $E(N;\lambda)$ when $\lambda_0 = 1.0$. The bounds to the continuation region depend only on the risks α and β and the ratio λ_1/λ_0 and will then be the same for all the tests above. The following is then obtained for the three truncated SPRT's

Period	Truncation point	Boundaries to continuation region	Average sampling number $\lambda = \lambda_0$	$\lambda = \lambda_1$
March-April	44	± 3.846	24	20
May-July	28	± 3.846	15	13
August-October	22	± 3.846	12	10

A decision procedure with three levels is used as for fatal accidents (see Section VII.2). At the time of the changeover, day, week and four-week tests were used. The construction of these tests was described in Section VII.2. All the tests were studied using data from the period Sept. 4 - Oct. 31, 1967. For this period both the day-by-day and the week-by-week test always gave the decision "no change". The results of the four-week test, the fixed-sample size test and the truncated SPRT are presented in Table VII.5. As before, it should be kept

in mind that the different tests are based on different statistical formalizations of the same problem. The fixed-sample test is constructed to provide a decision between H_0 and H_1, so that we do not obtain any result equal to "no change". As found in Chapter V, the truncated SPRT have lower actual risk levels than the intended ones.

Table VII.1 Truncation point (N_0) for one-sided and two-sided truncated SPRT for different values of α, β and λ_1 when choosing between $H_0: \lambda \leq 0.83$ and $H_1: \lambda \geq \lambda_1$

	$\lambda_1 = k \cdot \lambda_0$			
k	One-sided SPRT		Two-sided SPRT	
α β	1.25	1.50	1.25	1.50
0.05	297	75	292	80
0.01	598	151	585	162

Table VII.2 The average sampling number for one-sided and two-sided truncated SPRT when $\lambda = \lambda_1$ is true for different values of α, β and λ_1, when choosing between $H_0: \lambda \leq 1.0$ and $H_1: \lambda \geq \lambda_1$

	λ_1			
λ_1	One-sided SPRT		Two-sided SPRT	
α β	1.25	1.50	1.25	1.50
0.05	109	27	136	34
0.01	177	45	-	-

Table VII.3 Results of different decision rules applied to fatal accidents in congested areas between Sept. 4 - Dec. 24, 1967, and dates when the different tests are terminated. The decision rules are based on truncated SPRT, a fixed-sample size test and a four-week test.

Datum	Termination of	Resulting decision
Sept. 10	week test	"increased safety level"
Sept. 14	truncated SPRT	-"-
Sept. 25	truncated SPRT	-"-
Oct. 1	four-week test	-"-
Oct. 29	four-week test	-"-
Oct. 6	truncated SPRT	-"-
Nov. 18	fixed-sample size test	-"-
Nov. 25	truncated SPRT	-"-
Nov. 26	four-week test	-"-
Nov. 24	four-week test	"no change"

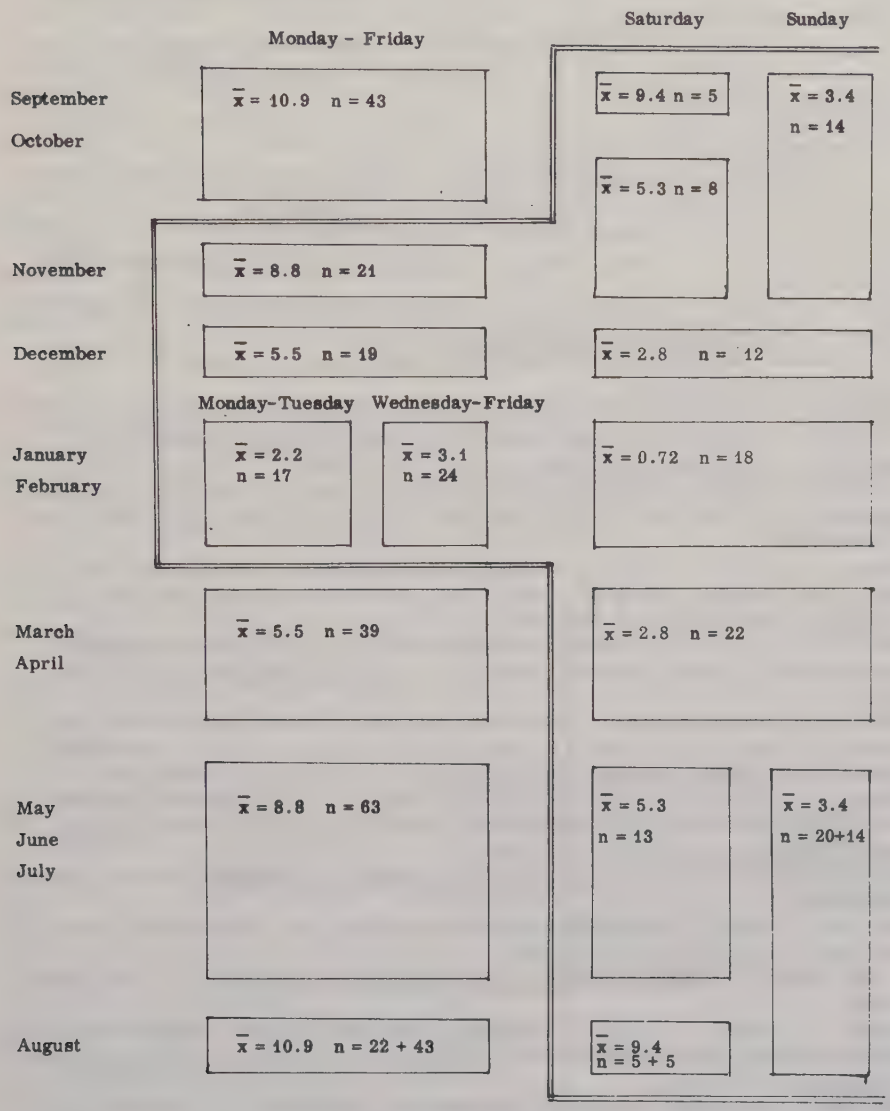
Table VII.4 Partition into groups of accidents with personal injury between motor vehicles and bicycles or mopeds using AID analysis

Month	Weekday	$\frac{\hat{V}(x) - \bar{x}}{\bar{x}}$	$\bar{x}$	$\hat{\delta} = \frac{\bar{x}^2}{\hat{V}(x) - \bar{x}}$	Number of observations
Jan., Feb.		0.85	2.2	2.6	59
	Monday, Tuesday	1.32	2.2	1.7	17
	Wednesday-Friday	0.77	3.1	4.0	22
	Saturday-Sunday	-	0.7	-	20
March, Apr., Dec.	Monday-Friday	0.05	5.5	108.0	62
May-July, Nov.	Monday-Friday	0.22	8.8	39.5	82
Aug.-Oct.	Monday-Friday	-0.02	10.9	-	66
March, Apr., Dec.	Saturday-Sunday	0.38	2.8	7.3	30
May-Nov.	Saturday	1.14	6.5	5.7	31
May-July, Oct.-Nov.	Saturday	1.94	5.3	2.7	24
Aug.-Sept.	Saturday	0.46	9.4	20.5	7
May-Nov.	Sunday	0.04	3.4	82.6	35

Table VII.5 Results of different decision rules applied to accidents with personal injury between vehicles and bicycles or mopeds in congested areas between Sept. 4 - Oct. 31, 1967, and dates when the different tests are terminated. The decision rules are based on truncated SPRT, a fixed-sample size test and a four-week test

Datum	Termination of	Resulting decision
Sept. 15	truncated SPRT	"decreased safety level"
Oct. 1	four-week test	-"-
Oct. 1	fixed-sample test	-"-
Oct. 3	truncated SPRT	-"-
Oct. 28	truncated SPRT	"no change"
Oct. 29	four-week test	-"-
Oct. 29	fixed-sample test	"increased safety level"

Figure VII.1 Partition into groups of accidents with personal injury between motor vehicles and bicycles or mopeds using AID analysis with number of observations (n) and average number of accidents per calendar day ($\bar{x}$) in each group



REFERENCES

Anderson, T.W. (1960). A Modification of the Sequential Probability Ratio Test to Reduce the Sample Size. The Annals of Mathematical Statistics, 31, 165-197.

Arbous, A.G., Kerrich, J.E. (1951). Accident Statistics and the Concept of Accident Proneness. Biometrics, 7, 340-432.

Armitage, P. (1957). Restricted Sequential Procedures. Biometrika, 44, 9-26.

Baker, A.G. (1950). Properties of Some Tests in Sequential Analysis. Biometrika, 37, 334-346.

Barraclough, E.D., Page, E.S. (1959). Tables for Wald Tests for the Mean of a Normal Distribution. Biometrika, 46, 169-177.

Bhate, D.H. (1959). Approximation to the Distribution of Sample Size for Sequential Tests. I. Tests of Simple Hypothesis. Biometrika, 46, 130-138.

Chanda, K.C. (1971). Asymptotic Distribution of the Sample Size for a Sequential Probability Ratio Test. Journal of the American Statistical Association, 66, 178-183.

Chatterji, S.D. (1963). Some Elementary Characterizations of the Poisson Distribution. American Mathematical Monthly, 70, 958-964.

Cochran, W.G. (1954). Some Methods for Strengthening the Common χ^2-test. Biometrics, 10, 417-451.

Dvoretzky, A., Kiefer, J., Wolfowitz, J. (1953). Sequential Decision Problems for Processes with Continuous Time Parameter. Testing Hypothesis. The Annals of Mathematical Statistics, 24, 254-264.

Elofsson, S.A. (1968). Uppföljning av olycksdata i samband med högertrafikomläggningen i Sverige 1967. Lic.avhandling. Institute of Statistics, University of Stockholm.

Elofsson, S.A. (1974). On the Use of an Asymptotic Distribution to the Sample Size of a Sequential Probability Ratio Test. Research Report. Institute of Statistics, University of Stockholm.

Erlander, S. (1970). A Review of Some Statistical Models Used in Automobile Insurance and in Road Accident Statistics. Acidental Analyses and Prevention, 3, 45-75.

Erlander, S., Gustavsson, J., Lårusson, E. (1968). Some Investigations on the Relationship between Road Accidents and Estimated Traffic. Research Report. Institute of Mathematical Statistics, University of Stockholm.

Erlander, S., Gustavsson, J., Svensson, Å. (1972). On Asymptotic Simultaneous Confidence Regions for Regression Planes in a Poisson Model. International Statistical Review, 40, 111-122.

Fisher, R.A. (1922). Statistical Methods for Research Workers. Olivier Boyd, Edinburgh.

Fisher, R.A. (1950). The Significance of Deviations from Expectations in a Poisson Serie. Biometrics, 6, 17-

Fisz, M. (1963). Probability Theory and Mathematical Statistics. Wiley, New York.

Freeman, D., Weiss, L. (1964). Sampling Plans which Approximately Minimize the Maximum Expected Sample Size. Journal of the American Statistical Association, 59, 67-88.

Gavatin-Avén, A. (1973). Signifikansproblemet i AID-analys. Lic.-avhandling. Institute of Statistics, University of Stockholm.

Gosh, B.K. (1970). Sequential Tests of Statistical Hypothesis. Addison-Wesley Publishing Company, Reading, Massachusetts.

Gustavsson, J. (1968). On the Use of Regression Models in the Study of Road Accidents. Research Report. Institute of Mathematical Statistics, University of Stockholm.

Gustavsson, J., Svensson, Å. (1972). An Application of a Poisson Regression Model to Classes of Road Accidents with Small Frequencies. Mimeographed Report. Institute of Mathematical Statistics, University of Stockholm

Hoel, D.G., Mazumdar, M. (1969). A Class of Sequential Tests for an Exponential Parameter. Journal of the American Statistical Association, 64, 1549-1559.

Johnson, N.L., Kotz, S. (1969). Distributions in Statistics: Discrete Distributions. Houghton Mifflin Company, Boston.

Kemp, K.W. (1958). Formulae for Calculating the Operating Characteristic and the Average Sampling Number of Some Sequential Tests. Journal of the Royal Statistical Society, Serie B, 20, 379-386.

Kendall, M.B., Stuart, A. (1967). The Advanced Theory of Statistics. Hafner, New York.

Kiefer, J., Weiss, L. (1957). Some Properties of Generalized Sequential Probability Ratio Tests. The Annals of Mathematical Statistics, 28, 57-74.

Lehman, E.L. (1953). Testing Statistical Hypotheses. Wiley, New York.

Miller, R.G. (1966). Simultaneous Statistical Inference. McGraw-Hill, New York.

Moran, P.A.P. (1952). A Characteristic Property of the Poisson Distribution. Proc. Cambridge Philosophical Society, 48, 206-207.

Morgan, J.N., Sonqvist, J.A. (1963). Problems in the Analysis of Survey Data — and a Proposal. Journal of the American Statistical Association, 58, 415-435.

Neyman, J. (1957). Current Problems of Mathematical Statistics. Proceedings of the International Congress of Mathematicians, 3, North Holland Publishing Company, Amsterdam.

Neyman, J., Pearson, E.S. (1928). On the Use and Interpretation of Certain Test Crituera for the Purpose of Statistical Inference. Biometrika, 20A, 175-240, 263-294. Also in Neyman & Pearson (1967). Joint Statistical Papers, 1-98. At the University Press, Cambridge.

Neyman, J., Pearson, E.S. (1933). On the Problem of the Most Efficient Tests of Statistical Hypotheses. Phil. Trans. Roy. Soc. A, 231, 289-337. Also in Neyman & Pearson (1967). Joint Statistical Papers, 140-185. At the University Press, Cambridge.

Page, E.S. (1954). An Improvement to Wald's Approximation for Some Properties of Sequential Tests. Journal of the Royal Statistical Society, Serie B, 16, 136-139.

Rao, C.R., Chakravarti, I.M. (1956). Some Small Sample Tests of Significance for a Poisson Distribution. Biometrics, 12, 264-282.

Roosmark, P.O., Fräki, R. (1969). Interview Investigation of Road Traffic Accidents. Accident Analysis and Prevention, 1, 279-292.

Sonqvist, J.A., Baker, E.L., Morgan, J.N. (1971). Searching for Structure. Institute for Social Research, University of Michigan, Ann Arbor, Michigan.

Sukhatme, P.V. (1938). On the Distribution of χ^2 in Samples of Poisson Series. Journal of the Royal Statistical Society Supplement, 5, 75-79.

Wald, A., (1945). Sequential Tests of Statistical Hypothesis. The Annals of Mathematical Statistics, 16, 117-186.

Wald, A., (1947). Sequential Analysis. Wiley, New York.

Weiss, L. (1953). Testing One Simple Hypothesis against Another. The Annals of Mathematical Statistics, 24, 273-281.

Weiss, L. (1962). On Sequential Tests which Minimize the Maximum Expected Sample Size. Journal of the American Statistical Association, 57, 551-566.

Wetherill, G.B. (1966). Sequential Methods in Statistics. Wiley, New York.

Zacks, S. (1971). The Theory of Statistical Inference. Wiley, New York.

Appendix A

Sequential probability ratio test of the mean in a Poisson distribution.

Study of the functions

$$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)} = h_p$$

$$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)} = h_N$$

$$h_{min} = \frac{\ln\left[\frac{\lambda_1 - \lambda_0}{\lambda \ln(\lambda_1/\lambda_0)}\right]}{\ln(\lambda_1/\lambda_0)}$$

used as approximations of the roots to the equation

$$\ln \varphi_\lambda(h) = 0 \ .$$

In order to study the three approximations of the roots to the equation $\ln \varphi(h) = 0$ given in Section II.2, we have to compare them with the value of h which gives the minimum of the function $\ln \varphi(h)$. This value, h_{min}, is given by

$$h_{min} = \frac{\ln\left[\frac{\lambda_1 - \lambda_0}{\lambda \ln(\lambda_1/\lambda_0)}\right]}{\ln(\lambda_1/\lambda_0)} .$$

Introduce the parameters Δ and k, given by

$$\Delta = \frac{\lambda_1 - \lambda_0}{\lambda_0} > 0 \; ; \quad (1 + k\Delta)\lambda_0 = \lambda . \tag{A.1}$$

Then the three approximations and h_{min} can be written as

$$h_A = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)} = \frac{-2[(1 + k\Delta)\ln(1 + \Delta) - \Delta]}{(1 + k\Delta) \cdot \ln^2(1 + \Delta)}$$

$$h_B = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)} = 1 - 2k \tag{A.2}$$

$$h_C = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)} = \frac{1 - 2k - k\Delta}{(1 + k\Delta)}$$

$$h_{min} = \frac{-\ln\left[\frac{(1 + k\Delta)\ln(1 + \Delta)}{\Delta}\right]}{\ln(1 + \Delta)} .$$

We now compare values of k for which the three functions h_A, h_B and h_C lie below and above the function h_{min}. Define $f_i = h_i - h_{min}$ for i = A, B and C and introduce

$$y = \frac{(1 + k\Delta)\ln(1 + \Delta)}{\Delta} . \tag{A.3}$$

We then have $\frac{\partial y}{\partial k} = \ln(1 + \Delta) > 0$ for all values of k.

The function $f_A(y)$ is

$$\begin{cases} f_A(y) = \dfrac{1}{\ln(1+\Delta)} \left\{ \ln y - \dfrac{2(y-1)}{y} \right\} \\ f'_A(y) = \dfrac{1}{\ln(1+\Delta)} \left\{ \dfrac{1}{y} - \dfrac{2}{y^2} \right\} \\ f''_A(y) = \dfrac{1}{\ln(1+\Delta)} \left\{ \dfrac{(4-y)}{y^3} \right\} \end{cases} \tag{A.4}$$

We now have

$$\begin{cases} f'_A(y) < 0 \text{ when } y < 2 \text{ and } f'_A(y) > 0 \text{ when } y > 2 \\ f_A(1) = 0 \qquad \lim_{y \to \infty} f_A(y) = +\infty \\ f_A(y) = 0 \text{ for some value } y_2 > 4.5 \ . \end{cases}$$

Then

$$\begin{cases} h_A > h_{min} \text{ for } y < 1 \\ h_A < h_{min} \text{ in any case for } 1 < y < y_1 > 4.5 \\ h_A > h_{min} \text{ for } y > y_1 > 4.5 \ . \end{cases}$$

The function f_B is

$$\begin{cases} f_B(y) = 1 - \dfrac{2y}{\ln(1+\Delta)} + \dfrac{2}{\Delta} + \dfrac{\ln y}{\ln(1+\Delta)} \\ f'_B(y) = \dfrac{1}{\ln(1+\Delta)} \left\{ \dfrac{1}{y} - 2 \right\} \\ f''_B(y) = - \dfrac{1}{y^2 \ln(1+\Delta)} \ . \end{cases} \tag{A.5}$$

We now have

$$\begin{cases} f_B'(y) > 0 \text{ for } y < \frac{1}{2} \; ; \quad f_B'(y) < 0 \text{ for } y > \frac{1}{2} \\ \lim_{y \to 0} f_B(y) = -\infty \quad f_B(\tfrac{1}{2}) = \dfrac{2+\Delta}{\Delta} - \dfrac{1 + \ln 2}{\ln(1+\Delta)} > 0 \quad \lim_{y \to \infty} f_B(y) = -\infty \\ f_B(y) = 0 \text{ for } y_1 < 1/4 \text{ and } 1 < y_2 < \dfrac{(1+\Delta)\ln(1+\Delta)}{\Delta} \; . \end{cases}$$

Then

$$\begin{cases} h_B < h_{min} \text{ for } y < y_1 < 1/4 \\ h_B > h_{min} \text{ for } 1/4 > y_1 < y < y_2 > 1 \\ h_B < h_{min} \text{ for } y > y_2 > 1 \; . \end{cases}$$

The function f_C is

$$\begin{cases} f_C(y) = \dfrac{2(1+\Delta)\ln(1+\Delta)}{\Delta^2 y} + \dfrac{\ln y}{\ln(1+\Delta)} - \dfrac{(2+\Delta)}{\Delta} \\ f_C'(y) = -\dfrac{2(1+\Delta)\ln(1+\Delta)}{\Delta^2 y^2} + \dfrac{1}{y \ln(1+\Delta)} \\ f_C''(y) = +\dfrac{4(1+\Delta)\ln(1+\Delta)}{\Delta^2 y^3} - \dfrac{2}{y^2 \ln(1+\Delta)} \; . \end{cases}$$

We now have

$$f_C' < 0 \text{ for } y < \frac{2(1+\Delta)\ln^2(1+\Delta)}{\Delta^2} \doteq y'$$

$$f_C'(y) < 0 \text{ for } y < \frac{2(1+\Delta)\ln^2(1+\Delta)}{\Delta^2} = y'$$

$$\lim_{y \to 0} f_C(y) = \infty \quad f_C(1) < 0 \quad f_C\{\frac{(1+\Delta)\ln(1+\Delta)}{\Delta}\} < 0 \quad \lim_{y \to \infty} f_C(y) = +\infty$$

$$f_C(y) = 0 \text{ for } y = y_1 \text{ and } y_2,$$

where

$$y_1 < 1 \text{ and } y_2 > \frac{2(1+\Delta)\ln^2(1+\Delta)}{\Delta^2} \; .$$

Then

$$h_C > h_{min} \text{ for } y < y_1, \text{ where } 1 < y_1 < \frac{(1+\Delta)\ln(1+\Delta)}{\Delta}$$

$$h_C < h_{min} \text{ for } y_1 < y < y_2, \text{ where } y_2 > \frac{2(1+\Delta)\ln^2(1+\Delta)}{\Delta^2}$$

$$h_C > h_{min} \text{ for } y > y_2 .$$

If the four functions are now compared in an interval $-1 < k < 2$, that is $0 < y < \frac{(1+2\Delta)\ln(1+\Delta)}{\Delta}$ we find

$$\begin{cases} y = 0 \qquad\qquad \lambda = 0 \\ \quad h_A \text{ and } h_C > h_{min}\ ,\ h_B < h_{min} \\ y_1 < 1/4 \text{ solution to } f_B(y) = 0, \text{ that is a value of } \lambda < \frac{\lambda_0}{2(2-\Delta)} \\ \quad h_A,\ h_B,\ h_C > h_{min} \\ y_2 \text{ solution to } f_C(y) = 0, \text{ where } y_2 > \frac{\Delta}{\ln(1+\Delta)}, \text{ that is } \lambda > \lambda_0 \\ \quad h_C < h_{min}\ ;\ h_A,\ h_B > h_{min} \\ y = 1 \qquad\qquad \lambda = \frac{\lambda_1 - \lambda_0}{\ln(\lambda_1/\lambda_0)} = \lambda' \\ \quad h_A,\ h_C < h_{min}\ ;\ h_B > h_{min} \\ y_3 \text{ solution to } f_B(y) = 0,\ y_3 < \frac{(1+\Delta)\ln(1+\Delta)}{\Delta}, \text{ that is } \lambda < \lambda' \\ \quad h_A,\ h_B,\ h_C < h_{min} . \end{cases}$$

For values of Δ smaller than or equal to 0.5, all other points discussed above are larger than $y = \frac{(1+2\Delta)\ln(1+\Delta)}{\Delta}$.

If we concentrate on values of Δ smaller than or equal to 0.5, the following is valid. All functions h_A, h_B, h_C lie on the same side of h_{min}, i.e. closer to the true value than h_{min},

at least in any case when $\lambda_0/3 < \lambda \leq \lambda_0$ and when $\lambda \geq \lambda_1$. In order to compare the results fully, values given in Appendix B also have to be taken into account (cf. Section II.2).

Appendix B

Sequential probability ratio test of the mean in a Poisson distribution.

Values of the functions

$$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$$

$$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$$

$$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)}$$

combined with values of $\ln \varphi_\lambda(h_0)$.

$\lambda_0 = 1.00 \qquad \lambda_1 = 1.10 \qquad \Delta = 0.10$

	$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
0.95	2.1913	0.0015	2.0000	-0.0005	2.1579	0.0011
0.96	1.9499	0.0011	1.8000	-0.0003	1.9167	0.0007
0.97	1.7135	0.0007	1.6000	-0.0002	1.6804	0.0004
0.98	1.4819	0.0005	1.4000	-0.0001	1.4490	0.0002
0.99	1.2549	0.0003	1.2000	-0.0000	1.2222	0.0001
1.00	1.0325	0.0002	1.0000	0.0	1.0000	0.0
1.01	0.8146	0.0001	0.8000	0.0000	0.7822	-0.0000
1.02	0.6009	0.0000	0.6000	0.0000	0.5686	-0.0001
1.03	0.3913	0.0000	0.4000	0.0000	0.3592	-0.0000
1.04	0.1858	0.0000	0.2000	0.0000	0.1539	-0.0000
1.05	-0.0158	0.0	0.0000	0.0000	-0.0476	0.0000
1.06	-0.2137	-0.0000	-0.2000	-0.0000	-0.2453	0.0000
1.07	-0.4078	-0.0000	-0.4000	-0.0000	-0.4392	0.0001
1.08	-0.5983	-0.0000	-0.6000	-0.0000	-0.6296	0.0001
1.09	-0.7853	-0.0001	-0.8000	-0.0000	-0.8165	0.0000
1.10	-0.9689	-0.0001	-1.0000	0.0	-1.0000	0.0
1.11	-1.1493	-0.0002	-1.2000	0.0000	-1.1802	-0.0001
1.12	-1.3264	-0.0004	-1.4000	0.0001	-1.3571	-0.0002
1.13	-1.5003	-0.0005	-1.6000	0.0002	-1.5310	-0.0003
1.14	-1.6712	-0.0007	-1.8000	0.0003	-1.7017	-0.0005

$\lambda_0 = 1.00 \qquad \lambda_1 = 1.25 \qquad \Delta = 0.25$

	$h_0 = \frac{-2\,[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
0.88	2.5132	0.0298	2.0000	-0.0078	2.4286	0.0222
0.90	2.1945	0.0200	1.8000	-0.0051	2.1111	0.0138
0.93	1.8929	0.0130	1.6000	-0.0031	1.8108	0.0079
0.95	1.6072	0.0080	1.4000	-0.0016	1.5263	0.0039
0.98	1.3362	0.0046	1.2000	-0.0006	1.2564	0.0014
1.00	1.0787	0.0025	1.0000	0.0	1.0000	0.0
1.02	0.8338	0.0012	0.8000	0.0003	0.7561	0.0006
1.05	0.6006	0.0004	0.6000	0.0004	0.5238	-0.0008
1.07	0.3782	0.0001	0.4000	0.0004	0.3023	-0.0006
1.10	0.1659	0.0000	0.2000	0.0002	0.0909	-0.0002
1.12	-0.0370	-0.0000	0.0000	0.0	-0.1111	0.0002
1.15	-0.2310	-0.0000	-0.2000	-0.0002	-0.3043	0.0006
1.17	-0.4168	-0.0002	-0.4000	-0.0003	-0.4894	0.0008
1.20	-0.5949	-0.0005	-0.6000	-0.0004	-0.6667	0.0008
1.22	-0.7656	-0-0010	-0.8000	-0.0003	-0.8367	0.0005
1.25	-0.9296	-0.0018	-1.0000	0.0	-1.0000	0.0
1.27	-1.0871	-0.0029	-1.2000	0.0005	-1.1569	-0.0009
1.30	-1.2386	-0.0043	-1.4000	0.0012	-1.3077	-0.0021
1.32	-1.3843	-0.0060	-1.6000	0.0022	-1.4528	-0.0037
1.35	-1.5246	-0.0082	-1.8000	0.0034	-1.5926	-0.0056

$\lambda_0 = 1.00 \qquad \lambda_1 = 1.50 \qquad \Delta = 0.50$

	$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
0.75	3.1776	0.3814	2.0000	-0.0625	3.0000	0.2812
0.80	2.6707	0.2272	1.8000	-0.0402	2.5000	0.1545
0.85	2.2235	0.1321	1.6000	-0.0238	2.0588	0.0793
0.90	1.8259	0.0740	1.4000	-0.0123	1.6667	0.0357
0.95	1.4702	0.0392	1.2000	-0.0046	1.3158	0.0118
1.00	1.1500	0.0191	1.0000	0.0	1.0000	0.0
1.05	0.8604	0.0081	0.8000	0.0023	0.7143	-0.0044
1.10	0.5971	0.0028	0.6000	0.0030	0.4545	-0.0047
1.15	0.3567	0.0006	0.4000	0.0025	0.2174	-0.0027
1.20	0.1363	0.0000	0.2000	0.0014	0.0000	0.0000
1.25	-0.0665	-0.0000	0.0000	0.0	-0.2000	0.0026
1.30	-0.2536	-0.0002	-0.2000	-0.0013	-0.3846	0.0046
1.35	-0.4269	-0.0011	-0.4000	-0.0021	-0.5556	0.0055
1.40	-0.5879	-0.0030	-0.6000	-0.0023	-0.7143	0.0051
1.45	-0.7377	-0.0060	-0.8000	-0.0017	-0.8621	0.0033
1.50	-0.8775	-0.0103	-1.0000	0.0	-1.0000	0.0
1.55	-1.0083	-0.0160	-1.2000	0.0028	-1.1290	-0.0048
1.60	-1.1309	-0.0230	-1.4000	0.0070	-1.2500	-0.0112
1.65	-1.2461	-0.0314	-1.6000	0.0125	-1.3636	-0.0190
1.70	-1.3546	-0.0411	-1.8000	0.0194	-1.4706	-0.0282

$\lambda_0 = 15.00 \qquad \lambda = 16.50 \qquad \Delta = 0.10$

	$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)\lambda}{\lambda(\lambda_1 - \lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
14.25	2.1915	0.0228	2.0000	-0.0075	2.1579	0.0171
14.40	1.9500	0.0162	1.8000	-0.0050	1.9167	0.0111
14.55	1.7136	0.0110	1.6000	-0.0031	1.6804	0.0067
14.70	1.4820	0.0072	1.4000	-0.0016	1.4490	0.0035
14.85	1.2551	0.0044	1.2000	-0.0006	1.2222	0.0013
15.00	1.0327	0.0024	1.0000	-0.0000	1.0000	-0.0000
15.15	0.8147	0.0012	0.8000	0.0003	0.7822	-0.0007
15.30	0.6010	0.0005	0.6000	0.0004	0.5686	-0.0009
15.45	0.3914	0.0001	0.4000	0.0004	0.3592	-0.0007
15.60	0.1859	0.0000	0.2000	0.0002	0.1539	-0.0003
15.75	-0.0158	0.0	0.0000	-0.0	-0.0476	0.0001
15.90	-0.2136	0.0000	-0.2000	-0.0002	-0.2453	0.0005
16.05	-0.4077	-0.0002	-0.4000	-0.0004	-0.4392	0.0008
16.20	-0.5982	-0.0005	-0.6000	-0.0004	-0.6296	0.0009
16.35	-0.7852	-0.0011	-0.8000	-0.0003	-0.8165	0.0006
16.50	-0.9689	-0.0021	-1.0000	0.0000	-1.0000	0.0000
16.65	-1.1492	-0.0035	-1.2000	0.0006	-1.1802	-0.0011
16.80	-1.3263	-0.0055	-1.4000	0.0014	-1.3571	-0.0027
16.95	-1.5002	-0.0080	-1.6000	0.0026	-1.5310	-0.0048
17.10	-1.6712	-0.0111	-1.8000	0.0042	-1.7017	-0.0076

$\lambda_0 = 15.00 \qquad \lambda_1 = 18.75 \qquad \Delta = 0.25$

	$h_0 = \frac{-2[\lambda \ln(\lambda_1/\lambda_0) - \lambda_1 + \lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1 + \lambda_0 - 2\lambda}{(\lambda_1 - \lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0 - (\lambda_1 + \lambda_0)}{\lambda(\lambda_1 - \lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
13.13	2.5132	0.4466	2.0000	-0.1172	2.4286	0.3337
13.50	2.1945	0.3000	1.8000	-0.0769	2.1111	0.2066
13.88	1.8929	0.1943	1.6000	-0.0465	1.8108	0.1180
14.25	1.6072	0.1201	1.4000	-0.0245	1.5263	0.0586
14.63	1.3362	0.0697	1.2000	-0.0094	1.2564	0.0212
15.00	1.0787	0.0371	1.0000	0.0	1.0000	0.0
15.37	0.8338	0.0173	0.8000	0.0049	0.7561	-0.0097
15.75	0.6006	0.0065	0.6000	0.0064	0.5238	-0.0114
16.12	0.3782	0.0017	0.4000	0.0055	0.3023	-0.0084
16.50	0.1659	0.0001	0.2000	0.0031	0.0909	-0.0028
16.87	-0.0370	0.0	0.0000	0.0000	-0.1111	0.0034
17.25	-0.2310	-0.0004	-0.2000	-0.0029	-0.3043	0.0087
17.62	-0.4168	-0.0023	-0.4000	-0.0050	-0.4894	0.0119
18.00	-0.5949	-0.0068	-0.6000	-0.0056	-0.6667	0.0119
18.37	-0.7656	-0.0146	-0.8000	-0.0041	-0.8367	0.0082
18.75	-0.9296	-0.0265	-1.0000	-0.0000	-1.0000	-0.0000
19.12	-1.0871	-0.0429	-1.2000	0.0072	-1.1569	-0.0130
19.50	-1.2386	-0.0641	-1.4000	0.0179	-1.3077	-0.0313
19.87	-1.3843	-0.0905	-1.6000	0.0326	-1.4528	-0.0550
20.25	-1.5246	-0.1223	-1.8000	0.0515	-1.5926	-0.0844

$\lambda_0 = 15.00$ $\lambda_1 = 22.50$ $\Delta = 0.50$

	$h_0 \frac{-2[\lambda \ln(\lambda_1/\lambda_0)-\lambda_1+\lambda_0]}{\lambda \ln^2(\lambda_1/\lambda_0)}$		$h_0 = \frac{\lambda_1+\lambda_0-2\lambda}{(\lambda_1-\lambda_0)}$		$h_0 = \frac{2\lambda_1\lambda_0-(\lambda_1+\lambda_0)\lambda}{\lambda(\lambda_1-\lambda_0)}$	
λ	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$	h_0	$\ln \varphi_\lambda(h_0)$
11.25	3.1776	5.7217	2.0000	-0.9375	3.0000	4.2187
12.00	2.6707	3.4077	1.8000	-0.6031	2.5000	2.3181
12.75	2.2235	1.9822	1.6000	-0.3575	2.0588	1.1888
13.50	1.8259	1.1105	1.4000	-0.1844	1.6667	0.5350
14.25	1.4702	0.5880	1.2000	-0.0694	1.3158	0.1764
15.00	1.1500	0.2860	1.0000	0.0	1.0000	0.0
15.75	0.8604	0.1219	0.8000	0.0348	0.7143	-0.0664
16.50	0.5971	0.0415	0.6000	0.0445	0.4546	-0.0698
17.25	0.3567	0.0090	0.4000	0.0374	0.2174	-0.0409
18.00	0.1363	0.0005	0.2000	0.0205	0.0000	0.0
18.75	-0.0665	-0.0001	0.0000	0.0000	-0.2000	0.0395
19.50	-0.2536	-0.0034	-0.2000	-0.0189	-0.3846	0.0689
20.25	-0.4269	-0.0168	-0.4000	-0.0318	-0.5556	0.0824
21.00	-0.5879	-0.0447	-0.6000	-0.0349	-0.7143	0.0767
21.75	-0.7377	-0.0902	-0.8000	-0.0252	-0.8621	0.0496
22.50	-0.8775	-0.1549	-1.0000	0.0000	-1.0000	0.0
23.25	-1.0083	-0.2398	-1.2000	0.0427	-1.1290	-0.0723
24.00	-1.1309	-0.3452	-1.4000	0.1045	-1.2500	-0.1674
24.75	-1.2461	-0.4711	-1.6000	0.1869	-1.3636	-0.2847
25.50	-1.3546	-0.6172	-1.8000	0.2907	-1.4706	-0.4236

Appendix C

Sequential probability ratio test of the mean in a continuous Poisson process.

Derivation of the first three moments of the decisive sampling time T from an exact expression of the moment generating function.

Let $x_1(t), t \geq 0$ and $x_2(t), t \geq 0$ be two different Poisson processes. Let H_i $(i=0,1)$ denote the hypothesis that $\lambda = \lambda_i$ with $\lambda_1 > \lambda_0$ and let T be the time when a decision is reached in a sequential probability ratio test. According to Dvoretzky et al.(1953) the moment generating function $M_\lambda(u,r) = E[e^{uT}]$ when the SPRT starts from r will be

$$M_\lambda(u,r) = \left(\frac{\lambda}{\lambda-u}\right)^{n(r)+1} + C(u)\cdot e^{-r(\lambda-u)/c} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!}\left[(J-r-i)\frac{\lambda}{c} e^{-(\lambda-u)/c}\right]^i -$$

$$- \frac{\lambda u}{(\lambda-u)^2} e^{(\lambda-u)(J-r-1)/c} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-(\lambda-u)/c}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{\lambda-u}{c}\right]^j \qquad (C.1)$$

where $c = (\lambda_1 - \lambda_0)/\ln(\lambda_1/\lambda_0)$, $n(r)$ is the integer such that

$$J-r-1 \leq n(r) < J-r, \quad K = b/\ln(\lambda_1/\lambda_0), \quad J = a/\ln(\lambda_1/\lambda_0)$$

and $C(u)$ is determined so that $M_\lambda(u,K) = 1$.

To determine the moments of T, we have to differentiate $M_\lambda(u,r)$ and then study the derivatives at $u = 0$. We then get

$$\frac{\partial M_\lambda(u,r)}{\partial u} = \frac{n(r)+1}{\lambda}\left(\frac{\lambda}{\lambda-u}\right)^{n(r)+2} +$$

$$+ \frac{K}{c} C(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!}[i+r]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ C'(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!}\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$- \frac{\lambda(J-r-1)u^2+\lambda u[c+\lambda(J-r-1)]+c\lambda^2}{c(\lambda-u)^3} e^{\frac{(\lambda-u)(J-r-1)}{c}} \times$$

$$\times \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{\lambda-u}{c}\right]^j - \frac{(c+\lambda-u)\lambda u}{c(\lambda-u)^3} e^{\frac{(\lambda-u)(J-r-1)}{c}}$$

$$\times \sum_{i=0}^{n(r)-1} i \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j +$$

$$+ \frac{\lambda u}{(\lambda-u)^3} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j .$$

Given $C(0) = 0$, we have

$$E(T) = \left.\frac{\partial M_\lambda(u,r)}{\partial u}\right|_{u=0} = \frac{n(r)+1}{\lambda} +$$

$$+ C'(0)\, e^{-\frac{\lambda r}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} \left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i +$$

$$- \frac{1}{\lambda} e^{(\lambda/c)(J-r-1)} \sum_{i=0}^{n(r)-1} e^{-(\lambda/c)i} \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{\lambda}{c}\right]^j$$

where $C'(0)$ should be determined so that

$$\left.\frac{\partial M_\lambda(u,K)}{\partial u}\right|_{u=0} \equiv 0 .$$

$$\frac{\partial^2 M_\lambda(u,r)}{\partial u^2} = \frac{[n(r)+1][n(r)+2]}{\lambda^2} \left(\frac{\lambda}{\lambda-u}\right)^{n(r)+3} +$$

$$+ \frac{1}{c^2} C(u)\, e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} \left[(i+r)^2\right]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ \frac{2}{c} C'(u)\, e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} \left[i+r\right]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$- \frac{\{\lambda(J-r-1)^2u^3+[2c\lambda(J-r-1)-2\lambda^2(J-r-1)^2]u^2+[\lambda^3(J-r-1)^2+2c^2\lambda]u+4c^2\lambda^2-2c\lambda^3(J-r-1)}{c^2(\lambda-u)^4} \times$$

$$\times e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{\lambda-u}{c}\right]^j$$

$$+ \left(\frac{2\lambda(J-r-1)u^3 - u^2[3c\lambda(J-r-1)+4\lambda^2(J-r-1)-2c\lambda]}{c^2(\lambda-u)^4} + \right.$$

$$\left. + \frac{u[2\lambda^3(J-r-1)+2c\lambda^2(J-r-1)-3c^2\lambda]+2c^2\lambda^2-2c\lambda^3}{c^2(\lambda-u)^4}\right) \times$$

$$\times e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} i \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j +$$

$$+ \frac{\{2\lambda(J-r-1)u^2+u[3c\lambda-2\lambda^2(J-r-1)]+2c\lambda^2\}}{c(\lambda-u)^4} e^{-\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \times$$

$$\times \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j + \frac{\lambda u(c+\lambda-u)\cdot 2}{c(\lambda-u)^4} \cdot e^{\frac{(\lambda-u)(J-r-1)}{c}}$$

$$\times \sum_{i=0}^{n(r)-1} i \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!} \left[(J-r-i-1)\frac{\lambda-u}{c}\right]^j +$$

$$- \frac{\lambda u(\lambda+c-u)^2}{c^2(\lambda-u)^4} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} i^2 \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j +$$

$$- \frac{\lambda u}{(\lambda-u)^4} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{j(-1)^j}{(j-1)!} \left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j$$

We then have

$$E(T^2) = \left.\frac{\partial^2 M_\lambda(u,r)}{\partial u^2}\right|_{u=0} = \frac{[n(r)+1][n(r)+2]}{\lambda^2} + \frac{2}{c} C'(0)\, e^{-\frac{r\lambda}{c}} \times$$

$$\times \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} (i+r)\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i + C''(0)\cdot e^{-\frac{\lambda r}{c}} \sum_{i=0}^{n(r)-1} \frac{(-1)^i}{i!}\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i +$$

$$- \frac{4c^2-2c\lambda(J-r-1)}{c^2\lambda^2} e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{j!} [(J-r-i-1)\frac{\lambda}{c}]^j +$$

$$- \frac{2}{\lambda^2}[1+\frac{\lambda}{c}] e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} ie^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{j!} [(J-r-i-1)\frac{\lambda}{c}]^j +$$

$$+ \frac{2}{\lambda^2} e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!} [(J-r-i-1)\frac{\lambda}{c}]^j$$

where C''(0) should be determined so that

$$\left.\frac{\partial^2 M_\lambda(u,K)}{\partial u^2}\right|_{u=0} \equiv 0$$

$$\frac{\partial^3 M_\lambda(u,r)}{\partial u^3} = \frac{[n(r)+1][n(r)+2][n(r)+3]}{\lambda^3}(\frac{\lambda}{\lambda-u})^{n(r)+4} +$$

$$+ \frac{1}{c^3} C(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} [(i+r)^3]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ \frac{3}{c^2} C'(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} [(i+r)^2]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ \frac{3}{c} C''(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} [(i+r)]\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ C'''(u) e^{-\frac{r(\lambda-u)}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} \left[(J-r-i)\frac{\lambda}{c} e^{-\frac{(\lambda-u)}{c}}\right]^i +$$

$$+ \frac{f_1(u)-3c\lambda^4(J-r-1)^2+12c^2\lambda^3(J-r-1)-18c^3\lambda^2}{c^3(\lambda-u)^5} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} \left(\frac{\lambda}{\lambda-u} e^{-\frac{\lambda-u}{c}}\right)^i$$

$$* \sum_{j=0}^{i} \frac{(-1)^j}{j!} [(J-r-i-1)\frac{\lambda-u}{c}]^j +$$

$$+\frac{f_2(u)+6c^2\lambda^3(J-r-1)+6c\lambda^4(J-r-1)-15c^3\lambda^2-12c^2\lambda^3}{c^3(\lambda-u)^5}\times$$

$$\times e^{\frac{(\lambda-u)(J-r-1)}{c}}\sum_{i=0}^{n(r)-1} i\left(\frac{\lambda}{(\lambda-u)}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j$$

$$+\frac{f_3(u)+15c^2\lambda^2-6c\lambda^3(J-r-1)}{c^2(\lambda-u)^5}e^{\frac{(\lambda-u)(J-r-1)}{c}}\times$$

$$\times\sum_{i=0}^{n(r)-1}\left(\frac{\lambda}{\lambda-u}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{(-1)^j}{(j-1)!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j+$$

$$+\frac{f_4(u)+6c\lambda^3+6c^2\lambda^2}{c^2(\lambda-u)^5}e^{\frac{(\lambda-u)(J-r-1)}{c}}\times$$

$$\times\sum_{i=0}^{n(r)-1} i\left(\frac{\lambda}{\lambda-u}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{(-1)^j}{(j-1)!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j+$$

$$+\frac{f_5(u)-3c\lambda^4-6c^2\lambda^3-3c^3\lambda^2}{c^3(\lambda-u)^5}e^{\frac{(\lambda-u)(J-r-1)}{c}}\times$$

$$\times\sum_{i=0}^{n(r)-1} i^2\left(\frac{\lambda}{\lambda-u}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j+$$

$$+\frac{f_6(u)-3c\lambda^2}{c(\lambda-u)^5}e^{\frac{(\lambda-u)(J-r-1)}{c}}\times$$

$$\times\sum_{i=0}^{n(r)-1}\left(\frac{\lambda}{\lambda-u}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{j(-1)^j}{(j-1)!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j+$$

$$+\frac{f_7(u)}{c(\lambda-u)^5}e^{\frac{(\lambda-u)(J-r-1)}{c}}\sum_{i=0}^{n(r)-1} i^2\left(\frac{\lambda}{\lambda-u}e^{-\frac{(\lambda-u)}{c}}\right)^i\sum_{j=0}^{i}\frac{(-1)^j}{(j-1)!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j+$$

$$+ \frac{f_8(u)}{c(\lambda-u)^5} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} i\left(\frac{\lambda}{\lambda-u} e^{-\frac{\lambda-u}{c}}\right)^i \sum_{j=0}^{i} \frac{j(-1)^j}{(j-1)!} \left[(J-r-i-1)\frac{(\lambda-u}{c}\right]^j +$$

$$+ \frac{f_9(u)}{c(\lambda-u)^5} e^{\frac{(\lambda-u)(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} i^3\left(\frac{\lambda}{\lambda-u} e^{-\frac{(\lambda-u)}{c}}\right)^i \sum_{j=0}^{i} \frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{(\lambda-u)}{c}\right]^j .$$

We then have

$$E(T^3) = \frac{[n(r)+1][n(r)+2][n(r)+3]}{\lambda^3} +$$

$$+ \frac{3}{c^2} C'(0)\, e^{-\frac{r\lambda}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} (i+r)^2\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i +$$

$$+ \frac{3}{c} C''(0)\, e^{-\frac{r\lambda}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!} (i+r)\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i +$$

$$+ C'''(0)\, e^{-\frac{\lambda r}{c}} \sum_{i=0}^{n(r)} \frac{(-1)^i}{i!}\left[(J-r-i)\frac{\lambda}{c} e^{-\frac{\lambda}{c}}\right]^i +$$

$$+ \frac{12c\lambda(J-r-1)-3\lambda^2(J-r-1)^2-18c^2}{c^2\lambda^3} e^{\frac{\lambda(J-r-1)}{c}} \times$$

$$\times \sum_{i=0}^{n(r)-1} e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{\lambda}{c}\right]^j +$$

$$+ \frac{6c\lambda(J-r-1)+6\lambda^2(J-r-1)-15c^2-12c\lambda}{c^2\lambda^3} e^{\frac{\lambda(J-r-1)}{c}} \times$$

$$\times \sum_{i=0}^{n(r)-1} ie^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{j!}\left[(J-r-i-1)\frac{\lambda}{c}\right]^j + \frac{15c-6\lambda(J-r-1)}{c\lambda^3} \times$$

$$\times e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!}\left[(J-r-i-1)\frac{\lambda}{c}\right]^j +$$

$$+ \frac{6\lambda+6c}{c\lambda^3} e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} ie^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{(j-1)!} [(J-r-i-1)\frac{\lambda}{c}]^j +$$

$$- \frac{3\lambda^2+6c\lambda+3c^2}{c^2\lambda^3} e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} i^2 e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{(-1)^j}{j!} [(J-r-i-1)\frac{\lambda}{c}]^j +$$

$$- \frac{3}{\lambda^3} e^{\frac{\lambda(J-r-1)}{c}} \sum_{i=0}^{n(r)-1} e^{-\frac{\lambda i}{c}} \sum_{j=0}^{i} \frac{j(-1)^j}{(j-1)!} [(J-r-i-1)\frac{\lambda}{c}]^j$$

where $C'''(0)$ should be determined so that

$$\left.\frac{\partial^3 M_\lambda(u,K)}{\partial u^3}\right|_{u=0} \equiv 0 .$$

Appendix D

Sequential probility ratio test of the mean in a Poisson distribution.

Comparison between numerical values of the first four moments of the decisive sampling number, N, derived from an approximate expression of the characteristic function of N, an exact expression of the moment-generating function of T and estimated from simulations of SPRT.

For the central moments of the decisive sampling number, N, derived from an approximate expression of the characteristic function of N and the decisive sampling time, T, λ_0 is a scale factor. This means that for given values of λ_1/λ_0, the moments of N and of T for λ_0 are given by

$$\frac{E\left\{(N - \mu_N)^k \middle|_{\lambda_0=1.0}\right\}}{\lambda_0^k} \quad \text{and} \quad \frac{E\left\{(T - \mu_T)^k \middle|_{\lambda_0=1.0}\right\}}{\lambda_0^k},$$

respectively.

The calculations of the moments of N and T in both of these cases are sensitive to the numerical precision used in the calculations. The moments of T are calculated with extended precision - given about 35 relevant figures - but still do not give accurate values when λ_1/λ_0 equals 1.1. The moments of N, when λ equals λ_0 and λ_1, are found by double precision - about 17 relevant figures - when $\alpha = \beta = 0.05$, but when $\alpha = \beta = 0.01$ and $\alpha = 0.01$, $\beta = 0.05$ extended precision had to be used. At $\lambda = (\lambda_1 + \lambda_0)/2$ accurate values of E(N) were found by double precision, while accurate values of the other moments could not be found even by extended precision. Tables D.1.1 - 2 list estimates of the moments of N from simulation. Values for λ_0 = 1.0, 5.0 and 10.0 are given in these tables. Tables D.2 - 3 give the first three central moments of N - calculated from the approximation used in this study - and of T when λ_0 equals 1.0 for risk levels $\alpha = 0.01$, $\beta = 0.05$ and $\alpha = \beta = 0.01$. The fourth central moments of N for these risk levels calculated from an approximate expression of the characteristic function of N are given in Table D.4.

Table D.1.1 The mean and variance of the decisive sampling number, N, and the decisive sampling time, T, when $\alpha = \beta = 0.05$

λ_0	λ_1	λ	E(N)			V(N)		
			E(T)	Approx.	Simulation	V(T)	Approx.	Simulation
1.000	1.100	1.000		565.05			160823	
		1.050		908.94				
		1.100		547.38			160825	
1.000	1.250	1.000	99.33	98.67	103.81	4792	4692	4658
		1.125	159.05	154.62	163.49	16866		16611
		1.250	94.05	91.60	96.69	4850	4692	5179
1.000	1.500	1.000	28.37	28.03	31.11	368	355	412
		1.250	44.47	42.05	47.86	1318		1614
		1.500	25.70	24.49	27.67	377	355	408
5.000	5.500	5.000		113.01	117.01		6433	6062
		5.250		181.79	198.42			26009
		5.500		109.48	114.41		6433	6906
5.000	6.250	5.000	19.87	19.73	22.40	191.66	187.67	234
		5.625	31.81	30.92	38.10	674.66		935
		6.250	18.81	18.32	21.57	194.00	187.69	234
5.000	7.500	5.000	5.67	5.61	7.27	14.72	14.19	21.73
		6.250	8.89	8.41	12.23	52.74		94.26
		7.500	5.14	4.90	6.70	15.06	14.20	23.03
10.000	11.000	10.000		56.51	60.93		1608	1759
		10.500		90.88	102.87			6804
		11.000		54.74	60.34		1608	1760
10.000	12.500	10.000	9.93	9.87	11.71	47.91	46.92	59.34
		11.250	15.91	15.46	20.60	168.66		256.59
		12.500	9.41	9.16	11.53	48.50	46.92	63.44
10.000	15.000	10.000	2.84	2.80	3.95	3.68	3.55	5.91
		12.500	4.45	4.21	7.15	13.18		28.51
		15.000	2.57	2.45	3.65	3.77	3.55	5.69

Table D.1.2 Third and fourth central moments of the decisive sampling number, N, and the decisive sampling time, T, when $\alpha = \beta = 0.05$

λ_0	λ_1	λ	$E(T-\lambda_T)^3$	$E(N-\mu_N)^3$ Approx.	$E(N-\mu_N)^3$ Simulation	$E(N-\mu_N)^4$ Approx.	$E(N-\mu_N)^4$ Simulation
1.000	1.100	1.000		$1.22 \cdot 10^8$			
		1.050					
		1.100		$1.22 \cdot 10^8$			
1.000	1.250	1.000	$6.30 \cdot 10^5$	$6.07 \cdot 10^5$	$4.90 \cdot 10^5$	$1.88 \cdot 10^8$	$1.26 \cdot 10^8$
		1.125	$4.29 \cdot 10^6$		$3.17 \cdot 10^6$		$1.35 \cdot 10^9$
		1.250	$6.34 \cdot 10^5$	$6.07 \cdot 10^5$	$5.93 \cdot 10^5$	$1.88 \cdot 10^8$	$1.53 \cdot 10^8$
1.000	1.500	1.000	$1.35 \cdot 10^4$	12638	14424	$1.07 \cdot 10^6$	$1.19 \cdot 10^6$
		1.250	$9.37 \cdot 10^4$		$1.07 \cdot 10^5$		$1.44 \cdot 10^7$
		1.500	$1.37 \cdot 10^4$	12643	14758	$1.07 \cdot 10^6$	$1.27 \cdot 10^6$
5.000	5.500	5.000		$0.97 \cdot 10^6$	$6.89 \cdot 10^5$	$3.52 \cdot 10^8$	$2.07 \cdot 10^8$
		5.250			$6.88 \cdot 10^6$		$4.00 \cdot 10^9$
		5.500		$0.97 \cdot 10^6$	$1.01 \cdot 10^6$	$3.52 \cdot 10^8$	$3.76 \cdot 10^8$
5.000	6.250	5.000	5040	4859	6748	$3.00 \cdot 10^5$	$4.50 \cdot 10^5$
		5.625	34325		52360		$6.68 \cdot 10^6$
		6.250	5074	4859	6440	$3.00 \cdot 10^5$	$4.26 \cdot 10^5$
5.000	7.500	5.000	108	101	178	1717	3683
		6.250	750		1526		54736
		7.500	109	101	207	1718	4337
10.000	11.000	10.000		$1.22 \cdot 10^5$	$1.27 \cdot 10^5$	$2.20 \cdot 10^7$	$2.16 \cdot 10^7$
		10.500			$9.50 \cdot 10^5$		$3.08 \cdot 10^8$
		11.000		$1.22 \cdot 10^5$	$1.39 \cdot 10^5$	$2.20 \cdot 10^7$	$2.57 \cdot 10^7$
10.000	12.500	10.000	630	607	784	18750	24362
		11.250	4291		6813		$4.11 \cdot 10^5$
		12.500	634	607	843	18752	26849
10.000	15.000	10.000	13.50	12.64	26.83	107	295
		12.500	93.69		215.35		3831
		15.000	13.62	12.64	21.88	107	215

Table D.2 The three first central moments of the decisive sampling number, N, and the decisive sampling time, T, when $\alpha = 0.01$ and $\beta = 0.05$

λ_0	λ_1	λ	$E(N)=\mu_N$ Approx.	$E(T)=\mu_T$	$V(N)$ Approx.	$V(T)$	$E(N-\mu_N)^3$ Approx.	$E(T-\mu_T)^3$
1.000	1.100	1.000	621		$2.23\cdot10^5$		$2.20\cdot10^8$	
		1.050	1443					
		1.100	663		$2.91\cdot10^5$		$2.61\cdot10^8$	
1.000	1.250	1.000	108	108	6488	6542	$1.10\cdot10^6$	$1.12\cdot10^6$
		1.125	245	250		44546		$1.85\cdot10^7$
		1.250	144	147	8460	8645	$1.30\cdot10^6$	$1.34\cdot10^6$
1.000	1.500	1.000	30.79	30.87	490	498	$2.29\cdot10^4$	$2.36\cdot10^4$
		1.250	66.97	69.71		3435		$3.98\cdot10^5$
		1.500	38.60	39.83	638	663	$2.71\cdot10^4$	$2.85\cdot10^4$

Table D.3 The three first central moments of the decisive sampling number, N, and the decisive sampling time, T, when $\alpha = \beta = 0.01$

λ_0	λ_1	λ	$E(N)=\mu_N$ Approx.	$E(T)=\mu_T$	$V(N)$ Approx.	$V(T)$	$E(N-\mu_N)^3$ Approx.	$E(T-\mu_T)^3$
1.000	1.100	1.000	960		$3.60\cdot10^5$		$3.83\cdot10^8$	
		1.050	2214					
		1.100	930		$3.60\cdot10^5$		$3.83\cdot10^8$	
1.000	1.250	1.000	168	168	10499	10559	$1.91\cdot10^6$	$1.94\cdot10^6$
		1.125	376	383		98015		$6.01\cdot10^7$
		1.250	156	158	10500	10687	$1.91\cdot10^6$	$1.95\cdot10^6$
1.000	1.500	1.000	47.64	47.75	795	803	$3.99\cdot10^4$	$4.08\cdot10^4$
		1.250	101.96	106.06		7508		$1.27\cdot10^6$
		1.500	41.62	42.90	795	821	$4.00\ 10^4$	$4.14\cdot10^4$

Table D.4 The fourth moment $E(N-\mu_N)^4$ of the decisive sampling number, N, when $\alpha = 0.01$, $\beta = 0.05$ and $\alpha = \beta = 0.01$ calculated from an approximate expression of the characteristic function of N.

λ_0	λ_1	λ	$E(N-\mu_N)^4$	
			$\alpha=0.01$ $\beta=0.05$	$\alpha=\beta=0.01$
1.00	1.10	1.00	$4.85 \cdot 10^{11}$	$1.03 \cdot 10^{12}$
		1.10	$6.27 \cdot 10^{11}$	$1.03 \cdot 10^{12}$
1.00	1.25	1.00	$4.13 \cdot 10^{8}$	$8.82 \cdot 10^{8}$
		1.25	$5.33 \cdot 10^{8}$	$8.82 \cdot 10^{8}$
1.00	1.50	1.00	$2.37 \cdot 10^{6}$	$5.07 \cdot 10^{6}$
		1.50	$3.05 \cdot 10^{6}$	$5.07 \cdot 10^{6}$

Appendix E

Derivation of an approximate expression of the ASN-function of two-sided truncated SPRT.

In Section II.3. the p.m.f. of N is approximated by

$$P(n;\lambda) \approx \frac{e^{-n\eta^2/2}}{n^{3/2}\sqrt{2\pi}} \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k \lambda_{kj} e^{-\lambda_{kj}^2/2n} \quad \text{when } \eta \neq 0$$

where the parameters are also defined in Section II.3. For the distribution function of N, we have

$$F(n;\lambda) = P(N \leq n/\lambda) \approx \int_0^n P(u;\lambda)\, du \ .$$

We can now write

$$F(n;\lambda) \approx \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k e^{-\lambda_{kj}\eta} \int_0^n \frac{1}{\sqrt{2\pi}} \lambda_{kj} u^{-3/2} [\exp - \tfrac{1}{2}(\lambda_{kj} u^{-1/2} - \eta u^{1/2})^2] du \ .$$

Write $\lambda_{kj} u^{-3/2}$ in the integrand as

$$\tfrac{1}{2}[\lambda_{kj} u^{-3/2} + \eta u^{-1/2}] + \tfrac{1}{2}[\lambda_{kj} u^{-3/2} - \eta u^{-1/2}]$$

and define

$$y = -\lambda_{kj} u^{-1/2} + \eta u^{1/2} \quad \text{and} \quad z = -\lambda_{kj} u^{-1/2} - \eta u^{1/2} \ .$$

We then have

$$dy = 1/2(\lambda_{kj} u^{-3/2} + \eta u^{-1/2})du \ ; \ dz = \tfrac{1}{2}(\lambda_{kj} u^{-3/2} - \eta u^{-1/2})du$$

giving

$$F(n;\lambda) \approx \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k e^{-\lambda_{kj}\eta} \left[\int_{-\infty}^{y_0} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy + \int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} e^{2\lambda_{kj}\eta}\, dz \right]$$

where

$$y_0 = -\lambda_{kj} n^{-1/2} + \eta n^{1/2} \ ; \ z_0 = -\lambda_{kj} n^{-1/2} - \eta n^{1/2} \ .$$

Integrating, we get

$$F(n;\lambda) \approx \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k \, e^{-\lambda_{kj}\eta} \left\{ \phi\left(\eta\sqrt{n} - \lambda_{kj} \frac{1}{\sqrt{n}}\right) + \right.$$

$$\left. + e^{2\lambda_{kj}\eta} \, \phi\left(-\eta\sqrt{n} - \lambda_{kj} \frac{1}{\sqrt{n}}\right) \right\} .$$

Truncating the SPRT at N_0 gives

$$P(N = N_0|\lambda) \approx 1 - \int_0^{N_0} P(u;\lambda)du = 1 - F(N_0;\lambda) =$$

$$= 1 - \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k \, e^{-\lambda_{kj}\eta} \left[\phi\left(\eta\sqrt{N_0} - \lambda_{kj} \frac{1}{\sqrt{N_0}}\right) + \right.$$

$$\left. + e^{2\lambda_{kj}\eta} \, \phi\left(-\eta\sqrt{N_0} - \lambda_{kj} \frac{1}{\sqrt{N_0}}\right) \right] .$$

For the average sampling number, we have

$$E_{N_0}(N;\lambda) \approx \int_0^{N_0} up(u;\lambda)du + N_0 P(N = N_0|\lambda) \approx$$

$$\approx \sum_{k=1}^{4} \sum_{j=0}^{\infty} \delta_k \lambda_{kj} \, e^{-\eta\lambda_{kj}} \int_0^{N_0} \frac{1}{\sqrt{2\pi}} u^{-1/2} \exp\left[-\frac{1}{2}(\lambda_{kj}u^{-1/2} - \eta u^{1/2})^2\right]du +$$

$$+ N_0 \cdot P(N = N_0|\lambda) .$$

Now write $u^{-1/2}$ as

$$\frac{1}{\eta} \left\{ \frac{1}{2}[\eta u^{-1/2} + \lambda_{kj}u^{-3/2}] + \frac{1}{2}[\eta u^{-1/2} - \lambda_{kj}u^{-3/2}] \right\}$$

and make the transformations

$$v = \eta u^{1/2} - \lambda_{kj}u^{-1/2} \quad \text{and} \quad \omega = \eta u^{1/2} + \lambda_{kj}u^{-1/2} .$$

Then

$$E_{N_0}(N;\lambda) \approx N_0 + \sum_{k=1}^{4} \sum_{j=0}^{\infty} \left(\frac{\lambda_{kj}}{\eta} - N_0 \right) \delta_k \, e^{-\eta \lambda_{kj}} \, \phi\left(\eta \sqrt{N_0} - \frac{\lambda_{kj}}{\sqrt{N_0}} \right) -$$

$$- \sum_{k=1}^{4} \sum_{j=0}^{\infty} \left(\frac{\lambda_{kj}}{\eta} + N_0 \right) \delta_k \, e^{\eta \lambda_{kj}} \, \phi\left(-\eta \sqrt{N_0} - \frac{\lambda_{kj}}{\sqrt{N_0}} \right)$$

when $\eta \neq 0$.

Appendix F

Average sampling number for fixed-sample test, SPRT, one-sided truncated SPRT and two-sided truncated SPRT.

Table F.1.1

			Risk levels		Fixed-sample test	Two-sided truncated SPRT						Non-truncated	One-sided truncated SPRT		
						Truncation point equal to N_F			Truncation point equal to 1.25 N_F						
λ_0	λ_1	λ	α	β	N_F	a	N_0	E(N)	a	N_0	E(N)	E(N)	a	N_0	E(N)
1.000	1.100	1.000	0.05	0.05	1136	6.519	11.36	994.30	3.903	1420	757.19	565.05	2.944	1534	1485.40
		1.025						1079.92			957.48				1380.60
		1.050										908.94			
		1.075						1072.13			933.06				842.02
		1.100						958.96			703.30	574.38			561.16
		1.150						616.60			373.19				281.83
		1.200						398.65			238.70				180.07
1.000	1.250	1.000			194	5.717	193.7	166.74	3.846	242.2	134.12	98.67	2.944	246.3	239.04
		1.050						176.15			158.59				227.73
		1.100						183.57			174.95				204.95
		1.125										154.62			169.06
		1.150						182.45			171.44				
		1.200						170.17			146.43				126.90
		1.250						147.02			112.49	91.60			90.09
		1.400						74.48			50.14				38.40
		1.500						51.03			34.13				26.28
1.000	1.500	1.000			53	5.183	53.08	44.18	3.776	66.35	38.84	28.03	2.944	63.16	60.55
		1.100						47.11			44.06				57.95
		1.200						48.96			47.45				52.54
		1.250										42.05			
		1.300						48.24			45.69				43.59
		1.400						43.66			38.04				32.60
		1.500						35.76			28.19	24.49			22.74
		1.750						18.00			13.11				10.23
		2.000						10.51			7.66				5.97

Table F.2.1

			Risk levels		Fixed-sample test	Two-sided truncated SPRT						Non-truncated	One-sided truncated SPRT		
						Truncation point equal to N_F			Truncation point equal to 1.25 N_F						
λ_0	λ_1	λ	α	β	N_F	a	N_0	E(N)	a	N_0	E(N)	E(N)	a	N_0	E(N)
1.000	1.100	1.000	0.05	0.01	1654							622.70	2.986	2452	2360.63
		1.025													2135.19
		1.050										1843.27			
		1.075													1009.02
		1.100										855.35			591.47
		1.150													285.80
		1.200													182.60
1.000	1.250	1.000			283							109.19	2.986	393.6	379.64
		1.050													355.20
		1.100													303.53
		1.125										244.69			
		1.150													225.45
		1.200													147.61
		1.250										141.39			94.96
		1.400													38.94
		1.500													26.65

Table F.2.2

			Risk levels		Fixed-sample test	Two-sided truncated SPRT						Non-truncated	One-sided truncated SPRT		
						Truncation point equal to N_F			Truncation point equal to 1.25 N_F						
λ_0	λ_1	λ	α	β	N_F	a	N_0	E(N)	a	N_0	E(N)	E(N)	a	N_0	E(N)
1.000	1.500	1.000	0.05	0.01	78							31.17	2.986	99.35	96.09
		1.100													90.35
		1.200													77.95
		1.250										66.97			
		1.300													58.43
		1.400													38.07
		1.500										37.10			23.97
		1.750													10.38
		2.000													6.05

Table F.3.1

			Risk levels		Fixed-sample test	Two-sided truncated SPRT						Non-truncated	One-sided truncated SPRT		
						Truncation point equal to N_F			Truncation point equal to 1.25 N_F						
λ	λ_1	λ	α	β	N_F	a	N_0	E(N)	a	N_0	E(N)	E(N)	a	N_0	E(N)
1.000	1.100	1.000	0.01	0.01	2272	7.542	2272	1512.60	5.149	2840	1103.04	962.91	4.595	3094	3073.14
		1.025						1935.02			1636.69				2946.17
		1.050										2213.74			
		1.075						1903.60			1584.43				1569.25
		1.100						1409.25			1008.46	927.31			913.22
		1.150						721.94			492.86				437.87
		1.200						461.27			314.91				281.03
1.000	1.250	1.000			389	6.747	388.0	247.38	5.089	485.0	197.84	168.71	4.595	496.7	493.56
		1.050						302.45			264.09				481.29
		1.100						342.52			324.04				438.49
		1.125										376.03			
		1.150						338.75			316.53				344.78
		1.200						283.20			238.84				229.50
		1.250						207.43			160.37	154.47			146.61
		1.400						87.99			66.36				59.92
		1.500						60.23			45.42				41.02

Table F.3.2

						Two-sided truncated SPRT							One-sided truncated SPRT		
			Risk levels		Fixed-sample test	Truncation point equal to N_F			Truncation point equal to 1.25 N_F			Non-truncated			
λ_0	λ_1	λ	α	β	N_F	a	N_0	E(N)	a	N_0	E(N)	E(N)	a	N_0	E(N)
1.000	1.500	1.000	0.01	0.01	108	6.199	106.65	68.01	5.021	133.3	58.57	48.12	4.595	125.4	124.66
		1.100						80.62			74.29				121.87
		1.200						90.55			88.38				111.93
		1.250										101.96			
		1.300						88.38			84.63				89.03
		1.400						70.73			61.84				59.17
		1.500						48.87			40.02	40.99			37.00
		1.750						21.55			17.45				15.97
		2.000						12.57			10.18				9.32

ISBN 91–7222–130–5
Sundt & Co Stockholm